Mentalizing the Self in Adolescence and its Links with Schizotypal Trait Expression

George Salaminios

Thesis Submitted for the Degree of Doctor of Philosophy

University College London (UCL)

Psychoanalysis Unit

I, George Salaminios confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Contemporary research suggests that clinical psychosis is distally linked with schizotypal trait expression and more proximally with the breakdown of psychological processes pertaining to mentalizing. Although previous findings are suggestive of a relationship between trait-vulnerability for psychosis and mentalizing difficulties, they involve adult participants either within or beyond the critical period of illness onset. To date, little is known about the process of mentalizing during the critical developmental period of adolescence or its associations with schizotypal trait dimensions.

In a series of empirical studies, the current thesis used novel experimental tasks and self-report measures in samples of typically-developing young people to: (1) examine the nature of associations linking schizotypal trait dimensions in adolescence to disruptions in mentalizing processes involving both the understanding of the self and others; (2) further understand the processes that sustain self-awareness during adolescence by examining the effects that age, cognitive effort and emotional valence may exert on self- and reality-monitoring performance; and (3) prospectively assess the nature of the relation between mentalizing processes sustaining self- (self-monitoring) and other-awareness (ToM) from adolescence to young adulthood.

Overall, the findings of the current thesis provide novel data suggesting that he expression of schizotypal traits that impede interpersonal communication with others in adolescence are associated with difficulties in self and other understanding. Regarding the development of psychological processes sustaining self-awareness, current data suggest that although both self- and reality-monitoring abilities may be established in pre-adolescent development, reality-monitoring capacities for emotionally-charged material may undergo further elaboration from adolescence to young adulthood. In addition, the data of the current

thesis suggest that increased cognitive effort and emotional valence during memory encoding may respectively lead to self- and reality-monitoring confusions. Finally, the findings of the current thesis suggest that different types of self-monitoring misattributions in adolescence can prospectively predict specific patterns of ToM dysfunction at 5-year follow-up.

Impact Statement

An increasing body of research suggests that mentalizing difficulties akin to those observed in people suffering with psychosis can also be identified, albeit at an attenuated degree, among non-clinical individuals in the context of schizotypal trait expression, prior to the development of clinical symptoms, suggesting a pathway towards illness expression. Despite these findings little is known at present about the relationship between schizotypal trait dimensions and mentalizing during adolescence. This is important as adolescence represents a key developmental period, both for the maturation and elaboration of mentalizing processes, but also for the emergence of the earliest trait signs of emerging psychosis.

The work presented in the current thesis provides novel cross-sectional and longitudinal findings on the nature of associations linking schizotypal trait dimensions to disruptions in mentalizing processes during the critical developmental period of adolescence. In addition, the current thesis offers important data regarding the process of adolescent mentalizing, particularly as it pertains to the monitoring of self-generated mental events.

This is the first comprehensive examination of the links between schizotypal trait dimensions and mentalizing processes underpinning both self- and other-awareness in adolescence. From a conceptual standpoint, the inclusion of typically-developing young people in the current thesis can increase our knowledge regarding factors that contribute to disrupt key mentalizing functions while these are still in the process of development. Furthermore, given that schizotypal trait expression and mentalizing difficulties both constitute vulnerability factors for psychotic illnesses, investigating the nature of their associations during adolescence is expected to increase scientific knowledge regarding the earliest stages of a pathogenic developmental process that relates to the expression of clinical psychosis.

Given the increasing body of clinical research highlighting the importance of treating emerging psychosis at its earliest stages, the findings of the current thesis are expected to

provide the empirical rationale to inform the development of adapted mentalization-based preventative interventions to attenuate psychosis vulnerability in young people who present with schizotypal trait manifestations, prior to the development of clinical symptoms. Specifically, the data of the current thesis (1) highlight the importance of encompassing evaluations of mentalizing abilities in the clinical assessment of psychosis risk during adolescence; and (2) provide the empirical basis to support the preventative application of mentalization-based psychotherapies (MBT) to sustain mentalizing development in young people confronted with trait risk for psychosis.

.

Peer-reviewed publications and conference presentations arising from the current thesis to date

Salaminios, G. and Debbané, M. (2021) A Mentalization-Based Treatment Framework to Support the Recovery of the Self in Emerging Psychosis During Adolescence. In Hasson-Ohayon, I. and Lysaker, P. (Eds.) *The Recovery of the Self in Psychosis: Contributions from Metacognitive and Mentalization Based Oriented Psychotherapy.* Routledge

Salaminios, G., Morosan, L., Toffel, E., Tanzer, M., Eliez, S., Badoud, D., Armando, M., Debbané, M. (2020) Associations Between Schizotypal Personality Features, Mentalizing Difficulties and Thought Problems in a Sample of Community Adolescents. *Early Intervention in Psychiatry*

Salaminios, G., Morosan, L., Toffel, E., Tanzer, M., Eliez, S., Debbané, M. (2020) Self-Monitoring for Speech and its Links to Age, Cognitive Effort, Schizotypal Trait Expression and Impulsivity during Adolescence. *Cognitive Neuropsychiatry,* 25 (3), 215-230

Debbané M.*, **Salaminios G.***, Luyten P., Badoud D., Armando M., Solida-Tozzi A., Fonagy P., Brent B.K., (2016) Attachment, Neurobiology and Mentalizing Along the Psychosis Continuum. *Frontiers in Human Neuroscience. 10, 406*
*joint first authors

Salaminios G., (2018) Schizotypal Trait Expression and Mentalizing in Adolescence. Oral presentation given at the *11th International Conference of Early Intervention in Mental Health.* Boston, USA, 7-10 Oct 2018

Salaminios G. (2017) *Self-Monitoring Patterns Across Adolescence and their links to Schizotypal Trait Expression.* Oral presentation given at the *2nd Meeting of the International Consortium of Schizotypy Research.* Beijing, China, 19-20 June 2017

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor, Dr Martin Debbané, for his essential support throughout this thesis, but most importantly for his interest in helping me grow in ways that extend far beyond the completion of this PhD. It is when I look back at where I was in the beginning of our collaboration that I can fully appreciate the deep impact that his sensitive and non-intrusive presence has had on helping me develop skills and personal qualities that will support me throughout my clinical and research careers. Working with him over the past years has been a constant source of inspiration and motivation.

I would also like to thank my second, supervisor, Dr Patrick Luyten, first for giving me the opportunity to undertake a PhD in the Psychoanalysis Unit and also for his input and encouragement throughout this work.

Next I would like to thank a number of colleagues, current and former, that have directly or indirectly contributed to this work. First, I would like thank Deborah, Larisa and Elodie at the University of Geneva. Their input in collecting the data used in the thesis and their support in helping me navigate through them made the completion of my PhD possible. I would also like to thank Michal at UCL for her helpful comments of my work. Next I would like to thank my former colleagues at the Early Years Parenting Unit of the AFC - Katja, Joe, Gemma, Leonie and Nicola – working alongside them was a meaningful experience and a constant source of learning. I am also thankful to Alan Colam at the Camden and Islington Psychodynamic Psychotherapy service for giving me the opportunity to undertake psychotherapeutic work while working on my PhD and for supporting my goal to train in psychoanalysis. Also, I would like to thank Amanda, David and Moya from the Thursday supervision group for their helpful comments on my clinical work and for making this journey more enjoyable. I would also like to thank Dr Yannis Malliaris for helping me begin my

research career and for teaching me very early on about the dimensionality of severe mental illnesses.

Next I want to thank my close friends in Athens (Antonis, Petros, Myrto, Bitzos, Christina) and in London (Dimitris, Aris, Themis, Koulis) for being there in good times and for helping me in the difficult times of the PhD and of life. I would also like to thank the person that has been the closest to me over the past year, her presence supports me in every way.

I am also thankful to my sister Eva, for supporting me in the beginning of my life in London as a 17-year-old psychology student and helping me in the first steps of my professional endeavours. Seeing her go through her own journey of completing her PhD has acted as a source of inspiration for me.

I would like to express my deep gratitude to Mrs. Tatiana Gracheva, who has been there for me for the past 7 years and throughout my PhD journey. She has helped me feel freer, more present and creative in my life. The completion of this thesis is only but a small indication of this.

I would like to dedicate my PhD to two people. To my father Filoktitis for everything he has done to support my personal and professional goals, for always being there to "photograph" the important moments in my life, for knowing when I was going through difficult times (even before I knew myself) and for helping me overcome them. Finally, I dedicate my PhD to my mother Christina. Although she is not here to see me complete it and I cannot share this moment with her, the sense of knowing how proud and happy she would have been with this and every other achievement in my life is always with me and helps me in everything I do.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

1.1.Rationale for the current thesis and overview of the chapter

Given the opportunity, every one of us can experience transient reality distortions akin to those observed in psychosis. For instance, experimentally-induced sensory manipulation has been shown to generate both hallucinatory experiences and depersonalization phenomena in community samples (Caputo, 2010; Daniel & Mason, 2015; Tsakiris et al., 2010). Similarly, psychoactive drugs, such as amphetamines and LSD that create states of heightened sensory awareness, are known to produce transient psychotic symptoms (Ham et al., 2017). Turning to everyday experience, it is safe to assume that when presented with unfamiliar situations that generate anxiety (i.e. walking in a quiet street at night and seeing two people approaching), we are liable to engage in momentary paranoid thinking involving fear, suspiciousness and a mistrust of others (Freeman, 2007).

Sensory and affective experiences are constitutive of what we perceive as "real" and provide us with a first-person perspective of ourselves in the world. Thus, it is not surprising that when exposed to experimental conditions, or ambiguous situations that significantly alter the amount and novelty of sensory and affective signals, we become prone to experience difficulties in distinguishing between reality and imagination. Fortunately, we usually manage to recover from these experiences with relative ease. Situations that generate uncertainty and challenge our sense of reality trigger in us a search for meaning in an attempt to understand and regulate our self-experience. We may implicitly, or explicitly, ask ourselves; "Did I really hear someone calling my name, or am I just tired?" and "Are these people really following me, or are they just walking in the same direction as me?" In doing so, we turn to higher order cognitive processes (thinking about our thoughts and feelings) in an attempt to compensate for

disruptions in our sensory system (Fonagy & Bateman, 2016, Frith, 1992). By intuitively engaging in a reflective process about our own and other peoples' thoughts and feelings, we quickly become aware that the voices we heard were products of our own imagination and that we may have erroneously ascribed malevolent intentions and self-referential meaning to others' behaviours. In other words, we manage to recover a more coherent sense of ourselves, as well as of ourselves in relation to others around us.

This is relevant for our understanding of psychotic experiences, which are not phenomenologically defined only by hallucinations or delusions, but also by a more enduring disturbance in the sense of self that may render oneself and one's place in the world as unrecognisable (Hasson-Ohayon, Cheli and Lysaker, 2021; Stern, 1985). More specifically, the experience of psychosis fundamentally involves a profound breach in the self's ordinary "going-on-being" in the world, which can present in various forms, including a diminished sense of personal agency, a loss of temporal awareness and perhaps most importantly, an undermined ability to genuinely relate to and be affected by others (Lysaker et al., 2018b). This view is reflected in contemporary formulations from phenomenologically-oriented approaches, which suggest that self-alterations in psychosis occur at the *pre-reflective* more embodied level of self-awareness. More specifically, phenomenological approaches postulate the subjective experience of psychosis is characterised by phenomena pertaining to hyper-reflexivity (exaggerated attention to normally tacit phenomena pertaining to the self) and diminished self-affectation (diminution of the sense of personal agency), which together contribute to compromise one's "taken for granted" sense of being the subject of their own experience (Sass and Parnas, 2003).Indeed, individuals diagnosed with psychosis and people who are at high risk for the illness often report chronic and pervasive forms of self-alienation, loss of personal meaning and lack of social connectedness, despite symptomatic improvement following psychotherapeutic or pharmacological treatment (Hamm et al., 2018). This has led to an

increasing interest in exploring the mental functions that can support the recovery of a basic sense of self in individuals living with psychosis and in those at increased risk (Lysaker et al., 2018b).

While phenomenologically-oriented approaches have highlighted the role of disturbances in pre-reflective self-awareness in psychosis, a different strand of conceptual and empirical research has drawn attention on how disruptions in higher order *reflective* processes may themselves contribute in the alterations in the sense of self seen as characteristic of psychosis (Frith, 2004; Frith, 1992). The higher order psychological processes that help us sustain a coherent understanding of ourselves and others in the face of increasing complexity can, at least in part, come under the term *mentalizing* – that is the imaginative capacity to understand the intentional mental states driving one's own and others' behaviours (Fonagy & Target, 1996).

Mentalizing, represents a multifaceted construct that encompasses a number of processes involved in mental state understanding, such as theory of mind (ToM: intentions from others' behaviours), empathy (understanding and relating to the emotional states of others), mindfulness (emotional self-awareness), and self-monitoring (cognitive self-awareness) (Choi-Kain & Gunderson, 2008; Debbané et al., 2016; Fonagy & Luyten, 2009). Good mentalizing entails the integration of these processes to construe a sensitive understanding of the internal imaginative states (i.e. thoughts, feelings) underpinning one's own and others' behaviours, while at the same time acknowledging the ever-changing subjective nature of mental states (Fonagy et al., 2002; Fonagy & Luyten, 2009). As such, mentalizing enables us to form flexible and predictive representational models of human subjectivity in order to distinguish inner from external reality, sustain a coherent sense of self and agency, and attenuate experiences of confusion and distrust within interpersonal relationships (Fonagy et al., 2002; Fonagy & Target, 1996).

Disruptions in the capacity to utilize mental state information to understand oneself and others have consistently been identified in people with chronic psychotic experiences and in those presenting with first episode psychosis (FEP) (Brent & Fonagy, 2014;). Indeed in a meta-analytic study ranking the mean effect sizes of 22 common neurocognitive tests in people with schizophrenia, mentalizing impairment was ranked 4th (Heinrichs and Zakzanis, 1998). Poor mentalizing in these samples has been linked to the severity of psychotic symptoms (Corcoran and Frith, 2003 ; Versmissen et al., 2008 ; Mazza et al., 2001; Harrington et al., 2005), reduced illness insight (Lysaker et al., 2005; Bora et al., 2007) and greater social dysfunction (Fett et al., 2011). Furthermore, empirical evidence suggest that the mentalizing disruptions in psychosis are independent from more generalised neurocognitive deficits in the domains of attention, memory and executive control (Brüne, 2005; Harrington et al., 2005). Overall, these findings have led to novel attempts to adapt mentalization-based psychotherapies (MBT) for the treatment individuals diagnosed with clinical psychosis (Brent, 2009, Brent, 2015, Weijers et al., 2020).

Critically, another line of research suggests that mentalizing difficulties are already present, albeit at an attenuated degree, during the preclinical stages of emerging psychosis, among help-seeking individuals who manifest transient psychotic states (Kim et al., 2011; Boldrini et al., 2020; Bora and Pantelis., 2013; Piskulic et al., 2016; Thompson et al., 2011) and in community samples of people reporting non-clinical psychotic-like experiences (Langdon and Coltheart, 1999; Clemmensen et al., 2014). These findings indicate that mentalizing difficulties do not represent outcomes of clinical psychosis, but may instead constitute factors that contribute in the developmental unfolding of the illness during its preclinical stages of expression. Given that contemporary clinical research highlights the value of treating emerging psychosis at its earliest stages, therapies targeting mentalizing difficulties can be applied preventively to attenuate psychosis risk, prior to the onset of clinically-relevant

symptoms (Debbané, Salaminios et al., 2016). Critical to this is the understanding of the nature of associations linking mentalizing processes to the earliest signs of psychosis expression.

Schizotypal personality traits involve subtle abnormalities in the perceptual, cognitive, emotional and interpersonal domains (Fonseca-Pedrero et al., 2018; Raine, 2006) and are understood to reflect the earliest phenotypic manifestations of an underlying genetic vulnerability for schizophrenia (Kwapil and Barrantes-Vidal; 2015, Grant et al., 2015; Lenzenweger, 2015). Surprisingly, to date, only a small number of correlational studies in community samples have examined the links between schizotypal traits or non-clinical psychotic-like experiences with mentalizing difficulties. These have reported associations between non-clinical perceptual aberrations (e.g. hallucination- and delusion-like phenomena) and reduced performance in mentalizing tasks involving self and other understanding (Langdon & Coltheart, 1999; Pickup, 2006). Although these findings are suggestive of a relationship between mentalizing dysfunction and premorbid signs of psychosis risk, they involve adult participants either within or beyond the critical period of illness onset. From an early prevention clinical standpoint, important insights can be gained by examining mentalizing processes in the context of schizotypal trait dimensions during earlier developmental stages.

The present thesis aims to empirically examine mentalizing processes pertaining to self and other understanding in typically-developing adolescents, as well as their associations with schizotypal trait dimensions and psychosis risk. The thesis focuses on adolescence as a critical period for the developmental elaboration of mentalizing processes (Choudhury et al., 2006; Dumontheil et al., 2010; Vetter et al., 2013), as well as for the emergence of schizotypal trait manifestations (Debbané & Barrantes-Vidal, 2014; Gooding et al., 2005). To frame the subsequent discussion and empirical studies of the thesis, the current introductory chapter is divided into two parts.

The aim of the first part is to provide the empirical and clinical rationale supporting the investigation of psychotic phenomena and experiences in non-clinical adolescents (i.e. in young people who are not diagnosed with schizophrenia-spectrum illnesses). To this end, the following sections begin by offering an overview of the transient symptom-like clinical risk states and more enduring trait-like manifestations that proceed the onset of clinical psychosis, their relevance for the development of psychotic illnesses, as well as the approaches developed to study these. The case will be made that (a) clinical state manifestations of psychosis risk (i.e. attenuated psychotic symptoms) during late adolescence and young adulthood reflect phenomenological markers of an already established, yet attenuated, psychotic process; while (b) trait manifestations of psychosis risk (i.e. schizotypal traits), which commonly emerge during late childhood and early adolescence, also reflect an underlying personality basis that may contribute in shaping the psychosocial developmental context to potentiate the unfolding of psychosis vulnerability towards late adolescence and young adulthood.

From this conceptual basis, the second part of the chapter will explore the relevance of mentalizing across the psychosis continuum, from its premorbid and prodromal stages to its clinical expression. More specifically, the chapter will present evidence suggesting that disruptions in psychological processes pertaining to the understanding of the self (i.e. self/reality monitoring) and others (i.e. ToM), which are commonly observed in people suffering with clinical psychosis, are also associated with subclinical phenomena, both at the state and trait levels of expression. It will be argued that investigating the nature of associations between emerging schizotypal traits and mentalizing processes during adolescence can inform our understanding regarding the first stages of a pathogenic developmental process that may modulate the expression of clinically-relevant psychotic phenomena. More specifically, it will be suggested that the expression of schizotypal traits during the critical developmental period of adolescence may causally contribute to disrupt the normative development of mentalizing

processes, thus increasing the risk for the emergence of clinical manifestations. The review of the literature will highlight potential areas of empirical interest in the study of mentalizing and schizotypy, including the relative lack of studies in adolescent populations. The final section of the chapter will present an overview of the empirical studies conducted in the current thesis, the hypotheses tested, as well as the methodologies adopted to investigate these.

Part one

1.2. The psychosis continuum

Historically, biomedical research on the pathophysiology and treatment of psychosis has been guided by a dichotomous categorical approach that clearly demarcates "normality" from illness on the basis of well-established diagnostic criteria, assessed primarily within the domains of perception (positive symptoms: hallucinations, delusions, suspiciousness/persecution), emotion (negative symptoms: blunted affect, emotional withdrawal, social withdrawal) and cognition (conceptual disorganization: disorganised language and communication, disorganised behaviour) (van Os et al., 1999). While the categorical approach to psychosis has proven clinically valuable in terms of informing diagnostic and treatment practices, because of its increased emphasis on clinical populations, its capacity to elucidate the complex dimensional nature of the illness has remained limited. Indeed, most recent conceptualizations posit that psychotic experiences do not represent categorical entities, but are expressed along a continuum ranging from mostly benign or moderate cognitive-perceptual aberrations and interpersonal difficulties, to subthreshold psychotic symptoms of lesser severity and duration and finally to the severe reality distortions typically identified in people diagnosed with a psychotic illness (Verdoux and van Os, 2002; Johns & Van Os, 2001). Within this model, core psychotic manifestations such as hallucinations and delusions are not seen as inherently tied to the clinical expression of the illness, the latter being primarily determined by other symptom

related dimensions such as their frequency, duration, intrusiveness and level of insight (Van Os et al., 2009).

An important assumption guiding the dimensional continuum-based approach is that psychotic phenomena are not only exhibited by people diagnosed with the illness, but can also be observed among non-clinical individuals from the general population, albeit at an attenuated level. Indeed, while the lifetime prevalence rate of schizophrenia is commonly reported at around 1% (McGrath et al., 2008; Murray et al., 2003), large-scale population data suggest that psychotic phenomena are relatively prevalent in the general population (Verdoux and van Os, 2002). For instance, epidemiological data by Kendler et al., (1996) showed that 28.4% from a general population sample responded positively to at least one question assessing psychotic symptomatology. Tien et al (1991), reported lifetime prevalence rates between 10-15% for hallucinatory experiences not related to organic causes. Another study found prevalence rates of 10% and 5% for paranoid ideation and hallucinatory experiences respectively (Eaton et al., 1991). Similarly, results from a birth cohort study indicated that 20.1% and 13.2% of subjects respectively reported at least one delusional and one hallucinatory experience by 26 years of age (Poulton et al., 2000).

Overall, population-level studies appear to support the notion of continuity between non-clinical psychotic-like manifestations and clinical psychosis. Most importantly perhaps, studies that have prospectively explored the psychiatric outcomes of individuals who report non-clinical psychotic-like experiences suggest an increased risk for the development of clinical psychosis, or for an increase in the frequency and severity of these manifestations within the next 10 years (Chapman et al., 1994; Poulton et al., 2000). This line of prospective evidence explicitly designates a developmental dimension to psychosis, in which clinical symptoms are understood to be preceded by and emerge from more subtle manifestations

during the preclinical stages of the illness (Debbané et al., 2014). These can be categorised into (1) relatively stable *trait* abnormalities; and (2) more transient *state* manifestations.

1.3. High-risk in psychosis: the trait and state risk approaches

Most contemporary conceptualizations of psychosis focus on two putative risk stages where the level of pathophysiology phenomenologically augments from subtle premorbid personality manifestations to prodromal psychotic states (Debbané, Salaminios et al., 2016; Debbané et al., 2014). These are understood to proceed the first episode of psychosis (FEP), which marks the conversion point at which symptomatic expression will for the first time reach the severity, duration and functional impairment required for a formal diagnosis of a schizophrenia spectrum illness.

The premorbid risk stage can be best characterised by the concept of *schizotypy* (Meehl, 1962; Claridge, 1997; Kwapil and Barrantes-Vidal, 2014), which encompasses a wide range of non-clinical and relatively stable *trait* abnormalities manifesting primarily within the perceptual, interpersonal and neurocognitive domains. These typically emerge during pre-adolescent and adolescent development and are subjectively experienced as common aspects of one's personality functioning (Kwapil and Barrantes-Vidal, 2014). Following the premorbid period, the prodromal stage most proximally linked to clinical psychosis during late adolescence and young adulthood is best defined by the Clinical High Risk concept (CHR, Fusar-Poli et al., 2013), which captures newly emergent transient psychotic *states* signalling imminent risk for transition to clinical illness.

The next sections will review the approaches developed to assess the clinical risk state and trait-like personality abnormalities that respectively characterize the prodromal and premorbid periods of psychosis expression, as well as their predictive value for the emergence of psychotic disorders.

1.3.1. The state approach: Clinical High Risk for psychosis (CHR)

Contemporary research has put forward two approaches for the assessment of the symptomatic manifestations characterising the prodromal phase that most proximally precedes the onset of clinical psychosis (i.e. the first diagnosable psychotic episode). Both of these place particular emphasis on the detection of newly emergent risk *states* that phenomenologically resemble those of psychotic disorders, although they differ from the latter in terms of their severity, frequency and level of conviction.

First, the most commonly used ultra-high risk approach (UHR; Fusar-Poli et al., 2013) outlines three prodromal syndromes whose presence prospectively identifies individuals who are at imminent risk for the development of clinical schizophrenia. These three syndromes involve the presence of: (1) frank positive psychotic symptoms that are too brief or intermittent to constitute clinical psychosis (present for several minutes a day, 4 days per week for 1 month); (2) attenuated positive psychotic symptoms (i.e. positive psychotic symptoms of subthreshold frequency, or intensity with an average frequency of at least once per week in the past month); and (3) functional decline in the presence of genetic risk (i.e. functional decline in the context of having a first-degree relative with a diagnosis of a psychotic disorder or having oneself a diagnosis of schizotypal personality disorder) (Yung et al., 1998; Fusar-Poli et al., 2013). On the basis of these three syndromes, a number of interview-based measures have been developed to clinically diagnose individuals at clinical high risk for psychosis (McGlashan, 2001; Raballo et al., 2011). These measures are designed to evaluate newly emergent subclinical manifestations pertaining to unusual thought content, delusional ideas, suspicion, persecutory ideas, grandiosity perceptual abnormalities / hallucinations and disorganized communication. Their subclinical nature primarily refers to the transient nature of the symptoms, as well as the conserved ability of individuals who exhibit these manifestations to remain aware that these represent deviations from their normal day-to-day functioning. It must be noted that given their

newly emergent nature and marked impact on day-to-day functioning, UHR manifestations are commonly assessed among help-seeking samples, rather than individuals from the general population.

As it can be seen, the UHR criteria predominantly define the prodromal period of psychosis expression on the basis of positive psychotic states that have not reached clinical significance in terms of severity, duration and conviction. As such, because of their specificity in assessing psychotic states phenomenologically akin to those identified in clinical populations, UHR criteria and their assessment tools have proven sensitive in capturing proximal risk for conversion to clinical psychosis among help-seeking populations. Indeed, the transition rate to a first-episode of psychosis among individuals who meet UHR criteria has been reported to range from 18% at 6 months, 22% at one year, 29% at two years and 36% at three years follow-up (Fusar-Poli et al., 2013). Overall, meta-analytic analyses of prospective findings have reported UHR transition rates spanning from 8% to 54% within 1 - 2.5 years of follow-up (Nelson et al., 2016)

The second, less commonly used approach seeking to capture CHR states relates to the assessment of *basic symptoms*, which entail a newly-emergent subjectively experienced sense of disruption in one's cognitive and/or perceptual processes (Schultze-Lutter et al., 2016). The symptoms assessed within this approach are primarily cognitive and perceptual in nature, including but not restricted to thought interference, unstable ideas of reference, derealisation, visual or auditory echoes, attentional problems and language difficulties. It must be noted that it is necessary that these symptoms occur at least weekly during the last three months to be considered relevant for proximal risk for psychosis. Thus, although perceptual and cognitive aberrations akin to those captured by the basic symptom approach may be experienced intermittently by non-clinical individuals across the lifespan, it is only when individuals report

a newly emergent, weekly occurrence of these symptoms that they signify a clinical state risk for psychosis.

Overall, the assessment of basic symptoms alone has been shown to yield lower transition rates (7 - 19% over a period of 6 – 18 months) compared to the assessment of UHR criteria (Ruhrmann et al., 2010). However, research has shown that it is possible, and clinically useful, to concurrently assess both UHR and basic symptom state manifestations when evaluating proximal risk for psychosis. A study carried out by Schultze-Lutter, Klosterkötter & Ruhrmann (2014) examined the value of combining the assessment of both UHR and basic symptoms to predict cumulative conversion rates of participants followed for a period of forty-eight months after their initial presentation (Schultze-Lutter et al., 2014). For individuals who only presented with either UHR manifestations or basic symptoms, the authors reported a transition rate of around 30%. However, among those presenting both attenuated positive symptoms and cognitive basic symptoms, a 70% conversion rate to clinical psychosis was identified. It thus appears that concurrent assessment of UHR criterial and basic symptoms, may significantly aid the clinical evaluation of proximal risk for psychosis.

1.3.2. CHR states as phenotypic markers of an already established psychotic process

Overall, the CHR approach has been effectively utilised to chart the prodromal risk states most proximally and phenomenologically linked to a first diagnosable episode of psychosis. From an early intervention clinical standpoint, identifying emerging psychotic states during the prodromal stage of the illness can support the timely application of indicated treatments to attenuate the risk for conversion to clinical psychosis and reduce the functional disability associated with the early course of the illness (McGorry et al., 2007; Fusar-Poli et al., 2020; Armando, Hustebaut and Debbané, 2019).

It must be noted however that because of its emphasis on diagnostically identifying psychotic states that signal imminent risk for clinical psychosis (i.e. within the next 2-3 years), the application of the CHR approach is often restricted to samples that are already within the critical developmental period of illness onset (i.e. late adolescents and young adults). For this reason, the capacity of the CHR approach to elucidate the earliest premorbid signs of psychosis risk, which often emerge during early-to-middle adolescence, remains limited. This is important given that early detection and intervention clinical strategies are increasingly moving towards younger age groups (McGorry et al., 2007).

Most importantly perhaps, evidence suggest that even when individuals who meet the CHR criteria do not transition to a first episode of psychosis, they do not generally display favourable mental health outcomes. For example, prospective studies have shown that in 34-82% of non-converters UHR symptoms remain stable for a period of 1-3 years (Simon et al., 2011), 40% of non-converters show poor functional outcomes after 3 years (Carrión et al., 2013) and 75% of non-converters are subsequently diagnosed with anxiety, affective, or substance abuse disorders at one-year follow-up (Addington et al., 2011). These evidence suggest that individuals who exhibit state manifestations of psychosis risk, whether they convert to clinical psychosis or not, are already within a clinical trajectory often characterized by adverse symptomatic and functional outcomes. In this sense, CHR indicators appear to capture the earliest state markers of an already established, or at least partly established, pathophysiological psychotic process. From a clinical standpoint, elucidating the premorbid factors that potentiate the development of clinical risk states can support the application of early prevention treatment approaches aiming to inhibit the trajectory of emerging psychosis at its earliest stages. In the next sections of the current chapter it will be proposed that trait manifestations pertaining to schizotypy may reflect the underlying premorbid personality basis upon which clinical psychotic states unfold across development.

1.3.3. The trait approach: Schizotypy

Modern approaches to the study of the premorbid signs of psychosis expression in non-clinical populations primarily rely on the psychometric evaluation of *schizotypy*, which captures personality trait manifestations pertaining to relatively stable aberrations in perceptual, interpersonal and neurocognitive processes. Similarly to schizophrenia, schizotypy is considered to represent a heterogenous construct, as evidenced at the phenotypic level, with manifestations ranging from mild disruptions in thought content and interpersonal functioning (e.g. odd beliefs, mild social withdrawal) to more marked aberrations in perception and cognition (e.g. hallucinatory experiences, disorganised thinking). Furthermore, schizotypy appears to follow the multidimensional structure of schizophrenia, which is comprised of markedly separate dimensions that differ with each other across the aetiological, developmental and treatment-response levels (Kwapil and Barrantes-Vidal; 2015). Indeed, most psychometric studies examining the factorial structure of schizotypy in the general population suggest that it is comprised of three separate and heterogenous dimensions that can be loosely mapped to the positive, negative and disorganized symptom dimensions of psychotic illnesses (Kwapil and Barrantes-Vidal; 2015, Grant et al., 2015; Lenzenweger, 2015). First, the cognitive-perceptual dimension (i.e. positive schizotypy) captures the tendency to go through unusual cognitive and perceptual experiences, which include perceptual oddities ranging from illusions to hallucinations, as well as disruptions in thought content pertaining to magical ideation, superstitious beliefs, delusional ideation, suspiciousness and paranoia. The interpersonal dimension (i.e. negative schizotypy) primarily involves an inability or diminution to experience pleasure from physical and social stimuli and its manifestations include flattened affect, lack of close friends, social withdrawal and a general sense of disinterest in others and the world. Finally, the disorganization dimension (i.e. disorganized schizotypy) captures experiences of

disorganized thinking that range from mild disturbances in thinking and behaviour to formal thought disorder and disorganized actions (Kwapil and Barrantes-Vidal; 2015).

According to most conceptualizations, schizotypal trait manifestations reflect the phenotypic expression of an underlying hereditary predisposition to schizophrenia (Kwapil and Barrantes-Vidal; 2015, Grant et al., 2015; Lenzenweger, 2015). Developmentally, these often emerge early in life and become clinically relevant during adolescence (Debbané & Barrantes-Vidal, 2014; Gooding et al., 2005). Contrary to the clinical risk states captured by the CHR approaches, which are experienced by affected individuals as signalling a marked deviation from their day-to-day functioning, schizotypal traits are subjectively experienced as common and stable aspects of one's personality functioning. Therefore, the assessment on schizotypal traits differs from the assessment of CHR state indicators in that it does not rely on "clinical state requirements" such as the newly emergent onset/worsening of manifestations or the level of functional disability that they exert for the individual. Accordingly, while CHR symptoms are most typically assessed among help-seeking individuals, schizotypal manifestations are primarily assessed within community samples and non-affected relatives of individuals suffering with psychosis. Historically, two complementary approaches have been used to guide the investigation of schizotypal phenomena in the general population.

The first approach, takes pathology as the reference point and defines schizotypy as a pathological personality organization exhibited only by a small number of genetically-prone individuals (i.e. "true schizotypes") (Meehl, 1962; Lenzenweger, 2015). According to this model, all individuals regarded as true schizotypes on the basis of their genetic liability will invariably exhibit a schizotypal personality organization, but only a small portion of those will go on to develop the clinically-diagnosable symptoms, while the rest will either remain asymptomatic (i.e. compensated) or only display stable but non-clinical manifestations across their lifespan (Meehl, 1953). From a conceptual standpoint, it is important to note that although

this model entails phenotypic variation in the expression of schizotypal manifestations along a continuum (i.e. from less to more severe), it retains a clinically categorical/taxonic component as the entire continuum is placed within the realm of illness (Grant et al., 2018). Thus, a person can either be a "schizotype" or not, although within the group of schizotypes, there is proposed variation regarding the severity of schizotypal manifestations. In this sense, this model of schizotypy has been characterised as *quasi-dimensional* (Claridge, 1997), as it posits a clear boundary between a healthy personality organization and a schizotypal one, while the latter is construed along a dimensional spectrum of expression, from mild schizotypal manifestations to the most severe end of schizophrenic symptoms. Measures developed to assess schizotypal traits on the basis of this model, primarily include items with a clinical content that are less frequently endorsed in the general population and are aimed for the clinical detection of "true schizotypes" (Mason, 2015). Although the items of these scales attempt to capture trait-like personality manifestations, they often retain their symptom-like nature, taking either single symptom features (e.g. delusional ideation or hallucination-proneness) or diagnostic criteria, most notably those of schizotypal personality disorder, as their starting point (see table 1.1. for a summary description of such scales).

Scale Name	Content	Subscales/items
Magical Ideation Scale (Eckblad and Chapman, 1983)	Thought transmission, astrology, spirit influences, reincarnation, psychic energy transfer, secret messages	Single scale with 30 items
Perceptual Aberration Scale (Chapman et al., 1978)	Unclear body boundaries, unreality, estrangement, change in body parts or appearance	Single scale with 23 items
Physical Anhedonia Scale (Chapman et al., 1976)	Lack of pleasure from sensory sources such as eating, touch, vision or sound	Single scale with 40 items
Revised Social Anhedonia Scale (Eckblad et al., 1982)	Schizoid indifference to other people	Single scale with 40 items

STA Scale (Claridge and Brooks, 1984)	Based on DSM-III Schizotypal Personality Disorder description	Single scale with 37 items
Schizotypal Personality Questionnaire (SPQ, Raine, 1991)	Based on DSM-III Schizotypal Personality Disorder criteria	3 scales with 9 subscales: cognitive perceptual (unusual perceptual experiences, ideas of reference, suspiciousness, odd beliefs or magical thinking), interpersonal (social anxiety, constricted affect, lack of close friends) and disorganization (odd speech, odd beliefs)
Peters Delusional Inventory (Peters et al., 1999)	Delusional ideas rated for presence, appraisal of distress, preoccupation and conviction	10 different domains of delusional ideas. 40 items in total

Table 1.1. Summary description of selected schizotypy scales based on the clinical approach to measurement (adapted from Mason, 2015)

Contrary to the quasi-dimensional model, which conceptualises schizotypy as a pathological personality organization that is present only among a small portion of vulnerable individuals, more recent conceptualizations stemming from the field of personality and individual differences have proposed a *fully-dimensional* approach in the understanding of schizotypy (Claridge & Beech, 1995). According to this model, schizotypy represents a multidimensional trait organization that is continuously distributed across the general population at varying degrees (Claridge, 1997). Within this model, schizotypy reflects a range of stable personality traits that manifest as individual differences in people's perception, emotional processing, thinking style and interpersonal functioning (Grant et al., 2018). In the fully-dimensional model, normal variation of personality is taken as the starting point when considering the schizotypy spectrum. As such, schizotypal traits adopt a dual quality, reflecting both individual differences in one's adaptive personality organization, but also a varying vulnerability for the development of psychotic disorders (Claridge, 1997). Accordingly, measures developed on the basis of this model tend to conceptualize schizotypal traits in the same vein as more global personality dimensions (e.g. neuroticism), thus avoid items that are

of extremely high or low difficulty and their content often extends beyond strictly psychopathological manifestations to also capture more subtle phenomena that are normally distributed in the general population (e.g. Deja vu experiences, impulsivity, attention difficulties) (see table 1.2. for a summary description of such scales).

Scale Name	Content	Subscales/items
Community Assessment of Psychic Experiences (Stefanis et al., 2002)	Positive psychotic-like phenomena; lack of emotions, motivation and social interest; cognitive symptoms of depression	Positive (18 items); negative (14 items); depression (8 items)
Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE; Mason, Claridge and Jackson, 1995)	Perceptual aberrations, magical thinking, hallucinations, purposelessness, moodiness, social anxiety, poor attention and decision-making, independence, solitude, social and physical anhedonia, avoidance of intimacy, impulsivity, eccentric behaviour, lack of self-control	Unusual experiences (30 items); cognitive disorganization (24 items); introvertive anhedonia (27 items); impulsive nonconformity (23 items)
Psychoticism scale (Eysenck et al., 1985)	Aggressive, cold, egocentric, impulsive, antisocial, creative, unempathic, tough minded, impulsive	Single scale with 32 items
Aberrant Salience Inventory (Cicero et al., 2010)	Assignment of salience to otherwise innocuous stimuli (i.e. reporting enhanced sensory, cognitive or emotional perception of stimuli)	Single scale with 29 items

Table 1.2. Summary description of selected schizotypy scales based on personality/individual differences approach to measurement (adapted from Mason, 2015)

It is important to note that despite the significance of the quasi-dimensional and fully-dimensional models in terms of generating and testing hypotheses regarding structure of the schizotypy construct, as well as its role in the development of psychotic illnesses, both models and the measures developed on their basis, suffer from a number of empirical limitations. For instance, the postulation of the quasi-dimensional model of schizotypy and its assertion of a schizotypal personality organization as categorically different from a "healthy" one, were

originally based on the idea of a single major gene accounting for the emergence of both schizotypal traits and schizophrenia (Meehl, 1962). However, the single-gene hypothesis of schizophrenia has not received empirical support, with the overwhelming majority of current conceptualizations and empirical studies suggesting that schizophrenia is a complex polygenic disorder (with each individual gene allele accounting for only small effect sizes). In a similar vein, the heterogeneity of schizophrenic and schizotypal manifestations also points to their multifactorial genetic and non-genetic underpinnings, thus raising further questions regarding the taxonic nature of the quasi-dimensional model. In contrast, as mentioned above, rather than adhering to a categorical view of schizotypy, the fully-dimensional model conceptualises schizotypal traits as being continuously distributed across the general population at varying degrees. Importantly however, empirical studies that have investigated the latent, as opposed to the phenotypic or psychometric, correlates (e.g. endophenotypic markers) of schizotypy in general population samples have not found evidence for their continuous distribution (Lenzenweger, 2015). Furthermore, it must be noted that although, in line with the categorical model, the fully-dimensional model postulates that high expressions of schizotypy signify increased risk for the development of clinical psychosis, it does not adequately clarify the effects that low or moderate expressions of schizotypy may engender for individuals. As such, neither model appears to have fully captured the complex nature of the schizotypy construct.

Indeed, more recent conceptualizations have attempted to integrate emerging empirical evidence to increase our understanding of schizotypy, as well as its role in the pathophysiology of psychotic disorders. For example, Lenzenweger (2015) introduced an adaptation of the quasi-dimensional model to suggest that schizotypy represents the outcome of developmental interactions between aberrant synaptic connectivity patterns (caused by a confluence of multiple risk-inducing genes, each contributing a small effect) and a number of environmental risk factors. According to Lenzenweger (2015), once schizotypy is consolidated, it may interact

with other psychosocial risk factors and polygenic potentiators to determine the emergence of various pathological outcomes, including schizophrenia, schizotypal personality disorder, or other schizophrenia-related psychotic disorders (e.g. delusional disorder). Thus, although this model does not place emphasis on a single-gene taxonic aspect, it manages to retain its clinical nature by conceptualising schizotypy as the latent psychopathological outcome of developmental interactions between risk-inducing genes and other pathogenic environmental influences.

Kwapil and Barrantes-Vidal (2015) have recently attempted to further elucidate the relation of schizotypy with an a number of other clinical constructs including schizophrenia, CHR states, schizotypal personality disorder and psychotic-like experiences. Specifically, they postulate that all these different psychosis-related constructs represent more narrow expressions of the schizotypy continuum and that any phenotypic differences between these simply reflect differences in how schizotypy is symptomatically expressed. As such, within model, schizotypy is understood to reflect a latent vulnerability for schizophrenia-spectrum illnesses that can be expressed across a very broad range of psychosis-related personality, subclinical and clinical phenomenology.

Interestingly, Debbané and Barrantes-Vidal (2015) have attempted to place schizotypy within a developmental framework by drawing attention to empirical evidence suggesting that the expression of schizotypal traits during the key window of adolescence may contribute to alter the normative developmental trajectories of various biopsychosocial processes linked to clinical psychosis. Against this background, they have conceptualised schizotypy as an underlying personality basis characterised by a number of risk-inducing properties that not only reflects the earliest manifest sign of emerging psychosis, but also causally impacts to shape the developmental context that will further potentiate the unfolding of psychosis vulnerability during the critical period spanning from adolescence to young adulthood (Debbané et al.,

2014). As it will be discussed in more detail in chapter 1.3.5. the model proposed by Debbané and Barrantes-Vidal (2015) provides a useful developmental framework for studying the dynamic and complex nature of bidirectional associations that may take place between schizotypy and other psychosis-relevant risk factors during adolescence. Therefore, in line with this developmental model, the empirical studies of the current thesis will conceptualise schizotypy as an underlying personality basis, whose expression in the form of schizotypal personality traits may thwart the normative trajectory of various psychosocial processes (including mentalizing) to augment the risk for the emergence of clinical manifestations. Accordingly, the emergence of clinical risk states (captured by the CHR approaches) or clinical psychosis during late adolescence and young adulthood will be understood as reflecting the outcome of a series of developmental transactions leading to the exacerbation of an underlying schizotypal trait liability (Debbané et al., 2014).

Given that the Schizotypal Personality Questionnaire (SPQ) was used across the empirical studies of the current thesis, the next paragraphs will describe its properties and critically review its use as a measure of schizotypy in the general population. First it is important to note that the SPQ was developed based on the DSM-III-R clinical criteria for schizotypal personality disorders (SPD) (Mason, 2015). SPD prevalence rates range between 2-3%, suggesting that its prevalence is higher than that of schizophrenia (Cadenhead and Braff, 2002). Importantly, SPD and schizophrenia are thought to share similar pathophysiological mechanisms and it is generally agreed that SPD reflects a trait marker for the development of schizophrenia (Yung et al., 1996). Furthermore, it has been reported that the conversion rate from SPD to schizophrenia-spectrum illnesses is over 11%, which is comparable to the conversion rates observed among the relatives of patients with schizophrenia (Debbané et al., 2015).

Most psychometric analyses examining the factorial structure of the SPQ consistently identify three main dimensions: the cognitive-perceptual (hallucination and delusion-like phenomena, magical thinking), interpersonal (social anxiety, constricted affect, lack of close friends), and disorganization dimensions (odd behaviours, odd speech) (Fonseca-Pedrero et al., 2018; Raine, 2006). Importantly, a number of empirical limitations pertaining to the factor structure of the SPQ and its use as a multidimensional measure of schizotypy must be noted.

First, as mentioned above, the SPQ was originally developed to assess SPD symptoms rather than schizotypy per se. For instance, the interpersonal dimension of the SPQ does not assess theoretically important aspects of the schizotypy construct, such as social or physical anhedonia, but instead includes social anxiety, a feature that has been shown to strongly correlate with other personality dimensions that involve emotion dysregulation and distress, such as neuroticism (Gross et al., 2014). Indeed, while the negative scales of other schizotypy measures have shown modest correlations with neuroticism, the correlations between the interpersonal dimension of the SPQ has been found to range from moderate to high (Gross et al., 2014). In a similar vein, the disorganized dimension of the SPQ appears to capture aspects of oddness and eccentricity, rather than dimensional aspects of cognitive disorganization or formal thought disorder (Oezgen and Grant, 2018). Indeed, the disorganized scale of the SPQ has been shown to correlate strongly with its cognitive-perceptual scale suggesting that the two assess, at least in part, overlapping features, thus may not represent distinct dimensions (Gross et al., 2014). In addition, high intercorrelations between the three dimensions have also been commonly observed, thus raising questions about the SPQ as a multidimensional measure of schizotypy (Gross et al., 2014). For instance, given that the SPQ represents a clinical measure of SPD traits, the high intercorrelations between its scales suggest that it may tap into more global aspects of distress and as such share variance with neuroticism (Oezgen and Grant, 2018). Overall, while empirical data support that the cognitive-perceptual dimension of the

SPQ does capture unique aspects pertaining to positive schizotypy, the evidence for its interpersonal and disorganized dimensions as measures of negative and disorganized schizotypy remain limited.

Given its theoretical development, as well as empirical data pertaining to its factor structure, it appears that the SPQ may not directly measure schizotypy. Rather, the SPQ may capture the manifest outcomes of schizotypy, namely schizotypal personality symptoms. Indeed, according to Kwapil and Barrantes-Vidal (2015), schizotypy may be expressed across a broad range of clinical and non-clinical phenomenology, which includes schizotypal personality disorder, as well as schizotypal personality traits. In other words, the SPQ may constitute a proxy measure capturing pathological outcomes of schizotypy, rather than schizotypy itself. Nonetheless, evidence also suggest that schizotypal personality features, as assessed by the SPQ do capture stable psychosis-like trait manifestations (Fonseca-Pedrero et al., 2018). Furthermore, the primary aim of the empirical studies of the current thesis will be to investigate how more "pathological" aspects of schizotypal trait expression may relate to mentalizing disruptions in community adolescents, thus justifying the use of the SPQ. Finally, given that the SPQ is one of the most commonly used measures of schizotypal traits, it's use in the current thesis also allows for comparisons between current findings and those of previous studies.

1.3.4. Schizotypy as a distal trait marker of psychosis expression

A number of general population longitudinal studies with study intervals spanning from 5 to 50 years suggest that self-rated schizotypal features, pertaining to both the cognitive-perceptual (e.g. perceptual aberration) and interpersonal (e.g. social anhedonia) dimensions, relate to the subsequent development of psychotic disorders. For example, Chapman et al (1994) found that 5% of individuals from a general population university sample scoring high on trait measures

of perceptual aberration or magical ideation developed schizophrenia over a 10 year period. Furthermore, 40% of individuals who scored high on both magical ideation and social anhedonia developed schizophrenia over the same study interval. In the same sample Kwapil et al (2013) reported that 24% of high scorers on social anhedonia developed schizophrenia-spectrum illnesses by 30 years of age. Similarly, a prospective study from a large birth cohort (n = 4871) found that those who developed psychotic disorders after the age of 31 years had previously reported high scores on social anhedonia (Miettunen et al., 2011). Bogren et al. (2010) found that a "paranoid-schizotypal" rating in a clinical interview of schizoid- and schizophrenia-related personality features predicted the subsequent development of clinical psychosis in a general population sample. Finally, Gooding et al (2005) reported that high scorers on schizotypy measures of perceptual aberration and magical ideation (n = 59) and of social anhedonia (n = 32) were significantly more likely to develop schizophrenia-spectrum illnesses over a 5-year interval compared to a group of individuals scoring low on these measures (n = 44). Overall, these findings suggest that self-rated schizotypy is longitudinally related to the emergence of psychotic illnesses. Furthermore, they highlight that research on schizotypy and psychosis-risk should not only include psychotic-like cognitive-perceptual manifestations (i.e. positive schizotypy), but also encompass multidimensional assessments that include measures of interpersonal schizotypal manifestations (i.e. negative schizotypy). Surprisingly, the longitudinal effects of the disorganised dimension on the development of psychotic disorders have not been adequately studied, thus it is important to consider that some of the effects attributed to positive and negative schizotypy may be confounded by the presence of disorganised manifestations. Indeed, the expression of schizotypal traits pertaining to disorganization has been shown to impact the developmental trajectory of cognitive-perceptual manifestations in adolescence (Debbané et al., 2013).

In line with general population studies, schizotypal trait dimensions have also been found to be statistically significant predictors for the later development of psychotic disorders in samples of non-clinical children, adolescents and young adults who are at genetic risk for the illness (i.e. first and second degree relatives of patients with schizophrenia). For instance, Erlenmeyer-Kimling et al (1993) showed that physical anhedonia in a child sample (mean age = 9.5 years) was significantly related to the development of psychosis within the next 11 years. In a sample of young adults at genetic risk (mean age = 21.2 years), Johnstone et al (2005) reported that total schizotypy scores, as well as schizotypal features pertaining to oddness and social withdrawal predicted the development of schizophrenia over a 5 year period. Finally, multiple structural equation data from a sample of genetic high risk individuals spanning from childhood to young adulthood (mean age = 15.9 years) showed that schizotypy was the only variable that was directly predictive of the emergence of clinical psychosis over a period of 3.5 years among a number of genetic (degree of relatedness to family member with schizophrenia), cognitive (IQ, working memory) and environmental risk factors (cannabis abuse, history of obstetric complications, removal from family home, parental separation or loss of a parent) (Shah et al., 2012).

Overall, schizotypal dimensions appear to be significant predictors of conversion to clinical psychosis. However, the measurement of schizotypal traits in non-clinical populations tends to yield significantly lower predictive values when compared to the assessment of CHR states measured among help-seeking individuals (Debbané et al., 2014). Furthermore, although the concurrent assessment of schizotypal traits has been found to statistically increase the predictive value of CHR measures, from a clinical point of view it's utility in terms of discriminating between those who convert to clinical psychosis and those who do not has been found to be limited (Mason et al., 2004). For these reasons, early detection clinical strategies to date have favoured the assessment of CHR syndromes over schizotypal personality traits

(Debbané et al., 2014). Indeed, the CHR approaches described above were developed under a criterion validation framework with the aim to capture individuals who are at imminent risk for developing clinical psychosis (thus their effectiveness has been assessed on the basis of transition rates among help-seeking populations). In contrast, the operationalization of the schizotypy construct and its psychometric assessment (even within the more clinical quasi-dimensional model) has always been guided by the assumption that expression of schizotypal traits in the general population may confer risk for a number of different outcomes, which may include among others, transition to clinical illness.

It must be noted however that although CHR state indicators are more effective in assessing imminent risk to clinical psychosis, they tend to lose their predictive value when assessed more than 3 years before the onset of clinical psychosis (Debbané et al., 2014). In other words, as discussed earlier, clinical risk states assessed by the CHR criteria appear to capture the first stages of an already established, albeit attenuated, psychotic process that links closely, both temporally and phenotypically, to clinical illness. In contrast, prospective studies with long-term study intervals in child and adolescent samples have shown that schizotypal traits are most effective as *distal markers* of psychotic illness onset (Debbané et al., 2014).

Interestingly, as it will be shown in the next section, the expression of schizotypal traits during childhood and early adolescence has been associated with a number of genetic, environmental and interpersonal risk factors for clinical psychosis (Shah et al., 2012). In this context schizotypal manifestations appear to define an early personality profile characterised by a number of underlying neurobiological and psychosocial properties upon which clinical states, including those captured by the CHR criteria, may unfold across development. As such, investigating schizotypal phenomena and their associations with other biopsychosocial risk factors in non-clinical populations can better inform our understanding of the earliest-stages of

psychosis pathogenesis than research restricted to individuals who are at the endpoint of the psychosis-liability continuum (Verdoux and van Os, 2002).

1.3.5. Schizotypy as a causal developmental determinant of psychosis risk in adolescence

Empirical studies suggest that the expression of schizotypal traits in community samples is significantly associated with a number of neuro-physiological, behavioural, and environmental aetiopathogenic mechanisms and risk factors typically linked to both the development and maintenance of clinical psychotic states. For instance, research has shown that schizotypal trait manifestations assessed among non-clinical individuals are associated with cognitive difficulties (e.g. in attention, memory and language production), adverse interpersonal experiences (e.g. attachment adversity, bullying) and exposure to environmental stress factors (e.g. living in urban environments) akin to those observed in people suffering with clinical psychosis (Ettinger et al., 2015; Korver-Nieberg et al., 2013; Pedersen et al., 2001). In a similar vein, a variety of genes and genetic polymorphisms identified as implicated in the aetiology of schizophrenia (e.g. dopamine-related genes) have also been associated with the expression of schizotypal traits in non-clinical samples (Barrantes-Vidal, Grant and Kwapil, 2015).

Importantly, research suggests that the associations between schizotypal traits and biopsychosocial factors linked to the expression of clinical psychotic states are already evident during early development. For instance, cluster analytic studies suggest that subgroups of adolescents scoring high on measures of schizotypy perform more poorly on neurocognitive measures of general intelligence and verbal fluency, obtain higher teacher ratings in terms of behavioural problems and display significantly more abnormalities in endophenotypic neurological markers (i.e. neurological soft signs, dermatoglyphic abnormalities) compared to young people with low scores on schizotypy measures (Barrantes-Vidal et al., 2003). In the interpersonal domain, studies have shown associations between traumatic life events,

maladaptive attachment patterns and self-rated schizotypal manifestations during childhood. For example, Escher et al. (2002), found that 86.3% from a sample of children experiencing cognitive-perceptual aberrations pertaining to auditory verbal hallucinations also reported experiencing traumatic and stressful life events at the time of their onset. Similarly, studies have reported significant associations between attachment avoidance and positive schizotypal expression (Korver-Nieberg et al., 2013; Berry et al., 2007; MacBeth et al., 2008), as well as between negative schizotypy and both attachment avoidance and anxiety (Berry et al., 2007; Sheinbaum et al., 2013). Interestingly, data from child and adolescent relatives of people diagnosed with psychosis showed that a number of cognitive (e.g. IQ), genetic (e.g. degree of relatedness) and environmental (e.g. parental loss) risk factors for psychosis were only indirectly related to the emergence of the illness after a 3.5 year interval and that was through the mediational role of self-reported schizotypy (Shah et al., 2012).

Overall, a growing body of literature indicates that risk factors typically related to the expression of clinical psychosis can also be identified among children and young people scoring high on psychometrically defined measures of schizotypy. These findings suggest that the developmental pathogenic process linking schizotypal personality traits to biological, interpersonal and behavioural features relevant for psychosis begins long before the onset of the first diagnosable symptoms of the illness in adulthood and prior to the emergence of CHR states in late adolescence (Debbané and Barrantes-Vidal., 2014). When this is considered conjointly with evidence suggesting that schizotypy constitutes a distal predictor for the development of clinical psychosis (Debbané et al., 2014), it highlights its potential role as a developmental determinant of psychosis risk - a causal factor characterised by a number of risk-inducing properties that not only reflects the phenotypic outcome of an emerging psychotic pathogenic process, but may also contribute in shaping the developmental context that will

further potentiate the unfolding of psychosis vulnerability during the critical period spanning from adolescence to young adulthood (Debbané et al., 2014).

Indirect evidence for the developmental impact of schizotypal traits on psychosis risk can first be identified at the endophenotypic neuro-physiological level. Endophenotypes for schizophrenia consist of measurable aberrations in neurophysiological processes that are closely linked to risk-inducing genes and cannot be readily observed at the phenotypic level. These subtle abnormalities typically involve low-level sensory functions (e.g. sensory-motor gating; auditory and visual processing) that are known to mature prior to pubertal development and contribute in sustaining the elaboration of higher-order cognitive processing (Kantrowitz et al., 2014). Endophenotypic impairments in these basic sensory functions are increasingly linked with the expression of schizotypal traits in children and young people from the general population, as well as in familial-risk youths who carry the risk genotype (Hans et al., 2009; Barrantes-Vidal et al., 2003; Ettinger et al., 2015; Nelson et al., 2013). Interestingly however, schizophrenia endophenotypes do not typically correlate with state manifestations of psychosis and do not appear to directly contribute to its symptomatic expression (Gottesman & Gould, 2003; Greenwood et al., 2013; Braff et al., 2001). Rather, it has been proposed that they exert their pathogenic impact by disrupting the normative developmental trajectory of higher-order cognitive processes (i.e. mentalizing), which are in turn directly linked to the core symptomatic expressions of the illness (Debbané, Salaminios et al., 2016). This would suggest that children and young people who report schizotypal traits undergo aberrant developmental trajectories for higher-order psychological processes on the basis of these underlying neuro-physiological anomalies, thus increasing their vulnerability to clinical psychosis from adolescence to young adulthood.

More direct indications for the developmental impact of schizotypal traits on psychosis vulnerability can be observed in the interpersonal and psychosocial domains. Empirical

findings suggest that although schizotypal manifestations (e.g. non-clinical auditory hallucinations) are relatively common in childhood and adolescence and in most cases naturally disappear spontaneously, their interaction with other psychosocial risk factors may lead to their persistence or exacerbation further along development (Brink et al., 2020; Bartels-Velthuis et al., 2016; Dominguez et al., 2011). Indeed, bidirectional and transactional interactions between schizotypal features and psychosocial risk factors are clinically common during childhood and adolescence. For instance, in the context of school environments, it is not uncommon for young people who exhibit disorganized manifestations pertaining to odd speech and eccentric behaviour to be met by their peers with hostility or even go through experiences of bullying-victimization (Debbané and Barrantes-Vidal, 2014). These adverse interpersonal interactions can in turn lead to experiences of social anxiety, social withdrawal (interpersonal schizotypy) and increasing suspiciousness (positive schizotypy), which together may augment psychosis vulnerability during adolescence. In a similar vein, experiences of childhood trauma or early attachment adversity may lead young people to appraise their social environment as hostile and threatening, again leading to schizotypal experiences of suspiciousness and paranoia (Pickering et al., 2008). In order to cope with these experiences, young people may engage in maladaptive copping strategies known to increase psychosis risk (i.e. cannabis use) or spend more time isolated and away from potentially positive social interactions with peers and other trustworthy adults that could otherwise attenuate their growing perception of their environment as threatening and persecutory.

Thus, from a dynamic developmental psychopathology standpoint, psychosis vulnerability during childhood and adolescence can be seen as developing though a complex matrix of bidirectional associations between emerging schizotypal personality traits and risk factors expressed across the neurobiological, interpersonal (e.g. attachment trauma, cannabis use, bullying) and psychological domains (Fonseca-Pedrero et al., 2020). In this context, the

emergence of clinical risk states (as captured by the CHR approaches) during late adolescence and young adulthood can be understood as reflecting the outcome of a series of developmental transactions leading to the exacerbation of an underlying schizotypal trait liability that is distributed in the general population at varying degrees (Debbané et al., 2014). This is supported by recent data suggesting that CHR adolescents exhibit significantly more schizotypal trait features than young people with a non-psychotic personality disorder diagnosis (Boldrini et al., 2020b).

Despite these findings, the psychological mechanisms linking schizotypal personality traits to the expression of psychotic states remain poorly understood. In the second part of this chapter, it will be proposed that mentalizing processes represent key psychological factors that may relate to and interact with an underlying schizotypal "seed" or trait liability to modulate the expression of psychosis vulnerability during adolescence.

1.3.6. Summary of part one

The first part of the introductory chapter provided an overview of the characteristics and predictive value of the prodromal risk states and premorbid trait manifestations that precede the full-blown expression of clinical psychosis, as well the approaches developed to assess these. By reviewing empirical findings to date, two key assumptions were highlighted: (1) clinical risk states (i.e. attenuated psychotic symptoms), which commonly emerge during late adolescence and young adulthood are experienced as marked deviations of one's day-to-day functioning and represent early phenotypic markers of an already established pathogenic trajectory that is linked, both proximally and phenomenologically, to clinical psychosis; and (2) the expression of schizotypal traits in pre-adolescence and early adolescence reflects an underlying personality basis characterised by a number of risk-inducing properties that may

causally impact on the psychosocial developmental context to potentiate the distal unfolding of psychotic states during late adolescence and young adulthood.

Against this conceptual background, the second part of this chapter will review evidence suggesting that mentalizing processes pertaining to both the understanding of oneself and others represent key psychological factors whose atypical development relates to psychosis risk across both the state/symptom and trait/personality levels of expression. It will be argued that investigating early associations between mentalizing difficulties and schizotypal traits during adolescence may help elucidate the first stages of a developmental pathogenic process that modulates psychosis vulnerability across development.

Part two

1.4. Mentalizing in the psychosis continuum

As discussed in the beginning of this chapter, mentalizing refers to the set of mental activities involved in the imaginative capacity to understand the intentional mental states (i.e. thoughts, emotions) driving one's own and others' behaviours (Fonagy & Target, 1996). Mentalizing represents a complex psychological construct that has been operationalised as organized along different but interrelated dimensions (Fonagy and Luyten, 2009). For instance, at any given moment, the focus of mentalizing activity may take as its object oneself or other people (self/other dimension). In a similar vein, the content of mentalizing activity of both the self and others may focus on cognitive (e.g. thoughts and intentions) or affective (e.g. emotions) material at varying degrees (cognitive/affective dimension). Furthermore, mentalizing judgments can be based on evaluations pertaining to externally observed characteristics (i.e. facial gestures) or inferences about internal states (internal/external dimension). Finally, mental state inferences can either be fast, intuitive and reflexive, or more declarative, slow and controlled (automatic/controlled dimension). Within this framework, different types of higher

order processes are understood to reflect different facets of the wider mentalizing construct (Choi-Kain and Gunderson, 2008). For example, ToM inferences about other people's thoughts or intentions would overlap with the cognitive, other-oriented facet of mentalizing. Similarly, self-oriented "metacognitive" processes involving the ability to recognise, think about and reflect on one's own thoughts, experiences and beliefs would overlap with the self-oriented aspect of mentalizing (Debbané et al., 2016).

Therefore, it can be seen that the construct of mentalizing does not seek to replace other conceptually-related constructs or processes (e.g. ToM). Rather, in its contemporary definition (used throughout this thesis) mentalizing captures all the higher-order reflective abilities that entail inferences about one's own and other peoples' mental states. This has often led to the concept of mentalizing being criticised as too wide, heterogenous, difficult to adequately define and hard operationalise (Choi-Kain and Gunderson, 2008). Indeed, each of the different processes encompassed within the mentalizing construct, such as ToM, metacognition or mindfulness, have their own conceptual traditions, empirical bases and unique measurement tools (Lysaker et al., 2021). Importantly however, large scale comparative investigations that would empirically clarify the relationship between these overlapping processes are currently lacking. As such, the study of different higher order processes has often remained scattered across multiple domains of investigation. Therefore, it can be argued that placing them under an integrative framework can facilitate the generation and testing of hypotheses pertaining to their development, the factors that contribute to their impairment, as well as their impact in the emergence of psychopathological outcomes.

Difficulties in utilizing mental state information to understand oneself and others are increasingly regarded as key factors that relate to the reality distortions characterising psychotic phenomena (Brent & Fonagy, 2014; van Os et al., 2010; Frith, 1992; Allen et al., 2007). Indeed, at the phenotypic level, a number of core psychotic phenomena explicitly involve difficulties

in recognising and reflecting on one's own or others' mental events. For instance paranoid delusional ideation entails the misattribution of other people's intentions (Brent and Fonagy, 2014), while auditory hallucinations and delusions of alien control involve difficulties in recognising and monitoring one's own self-generated inner mental events (i.e. inner speech, bodily signals).

A meta-analysis of data from 29 studies with a total number of over 1500 participants found that people suffering with psychosis exhibit significantly worse mentalizing performance in ToM tasks compared to healthy controls with a large mean effect size ($d = -1.25$) (Sprong et al., 2007). Furthermore, as it will be discussed in more detail in section 1.5, studies have shown that poor ToM performance in psychosis is linked with the severity of psychotic symptomatology (Frith., 2004). In addition, a different strand of empirical research has shown that people suffering with psychosis also experience self-reflective difficulties in recognising themselves as the source of their own thoughts and actions (Frith, 1992). Indeed experimentally-assessed impairments in self-monitoring (the capacity to recognise and monitor the source of self-generated mental events)(Johnson et al., 1993) have been shown relate to core psychotic manifestations. For instance, meta-analytic data from 15 studies including a total sample size of 489 schizophrenia patients found that difficulties in monitoring the source of self-generated verbal material was significantly associated to hallucinatory phenomena with moderate-to-large effect sizes (Brookwell, Bental and Varese., 2013). It must be noted that although self-monitoring is not typically studied under the construct of mentalizing, it does involve self-reflective inferences about mental events. Indeed, self-monitoring has previously been described in the literature as a metacognitive processes (Dimaggio et al., 2009; Brookwell, Bentall and Varese, 2013; Lysaker et al., 2011; Debbané et al., 2016; Debanné, Salaminios et al., 2016). As such, self-monitoring will be conceptualised throughout this thesis as a higher-order cognitive process that overlaps with the self-oriented cognitive facet of mentalizing.

Importantly, self-reflective difficulties in psychosis have often been assessed through narrative methodologies that aim to capture more integrative aspects pertaining to the capacity to understand and synthesise one's own personal experiences (Lysaker et al., 2005), or self-report measures that capture maladaptive beliefs about the process and content of one's thoughts (Varese and Bentall, 2011). Contrary to these, the experimental assessment of self-monitoring offers the opportunity to explore more fundamental disruptions in self-reflective abilities among people suffering with the illness and those at increased risk.

However, despite emerging empirical evidence and clinical conceptualizations regarding the links between mentalizing and psychosis, a number of gaps remain. First of all, although mentalizing impairments in ToM and self-monitoring appear to account for unique variance in symptomatic and functional outcomes in psychosis, it remains unclear how these contribute to the illness over and above other neurocognitive difficulties often identified in people suffering from psychosis. Indeed, a multiplicity of deficits in cognitive domains, such as working memory, attention and executive functioning have consistently been observed in psychosis (Heinrichs and Zakzanis, 1998). Furthermore, while most studies testing mentalizing abilities, such as ToM and self-monitoring often control for the effects of other neurocognitive impairments, it is likely that any test of ToM or self-monitoring would probably tap into multiple component processes at the same time. For example, tasks assessing ToM through the understanding of indirect speech would implicitly also rely on verbal comprehension abilities (Sprong et al., 2007). Thus, it remains possible that general cognitive abilities may represent a necessary, but perhaps not sufficient condition, for normative performance in mentalizing tasks (Sprong et al., 2007) Second, cognitive models do not explicitly clarify how impairments in higher order reflective processes such as self-monitoring and ToM relate to disturbances in pre-reflective self-awareness. For instance, it remains unclear whether impairments in reflective aspects pertaining to self- and other-understanding in psychosis occur independently from

disturbances in pre-reflective self-awareness, or whether they represent phenotypic correlates or even outcomes of the latter. Furthermore, within cognitive models of psychosis, disruptions in higher order reflective processes, such as self-monitoring and ToM are often viewed as "static" neurocognitive deficits, thus are seldomly placed within a developmental framework that would facilitate the study of the factors that contribute to impair these across development or for potential interconnections with pre-reflective aspects of the self. Most importantly perhaps, mentalizing dysfunction has been described as transdiagnostic and appears to be implicated in a number of psychopathological outcomes, including borderline personality disorder (Bateman & Fonagy, 2010); depression (Taubner et al., 2011) and functional somatic disorders (Luyten et al., 2012), thus raising questions pertaining to its specific role in psychotic disorders.

Interestingly, it has been suggested that common element that unifies the higher order functions that are subsumed under the mentalizing construct is that they all exert some form of top-down control to account for disruptions in lower-level information processing (e.g. sensory aberrations, emotion dysregulation) (Rudrauf, 2014; Fonagy and Bateman, 2016). This brings up the interesting possibility that disruptions in mentalizing abilities may not be causally associated to psychosis, but rather reflect the breakdown of resilience factors that could otherwise protect against transition to clinical illness among those who are at increased neurogenetic risk (Debbané; Salaminios, et al., 2016). Indeed, the underlying causes of psychotic illnesses can be traced to genetically-based disruptions in the detection, integration and processing of multi-sensory input (Greenwood et al., 2013). These can lead to alterations in pre-reflective self-awareness (characteristic of those described by phenomenological approaches) during the prodromal stages of the illness, which may involve aberrant sensory and perceptual experiences, such as derealization, depersonalization and hallucination-like phenomena (Postmes et al., 2014; Sass and Parnas, 2003). Importantly, failures to engage in a

reflective process or construe mental representations of these aberrant sensory and perceptual states may contribute to their exacerbation and increase the risk for the emergence of clinical symptoms. In other words, and from a clinical point of view, the critical value of mentalizing in emerging psychosis may be to afford resilience and protect against the development of the illness in the face of adverse sensory conditions (Salaminios and Debbané, 2021). While the role of mentalizing abilities as resilience factors in psychosis remains an empirically unexplored hypothesis, a few studies have shown that better mentalizing skills are linked to better outcomes during the clinical and preclinical stages of the illness. Indeed, Braehler and Schwannauer (2012) reported that better mentalizing capacities, assessed using narrative methodologies, supported adaptation and individuation processes in adolescents recovering from a first episode of psychosis. Furthermore, Bartels-Velthuis et al. (2011), reported that in sample of young adolescents who experienced non-clinical hallucinations within the previous five years, the risk for the development of secondary delusional ideation was lower for those with higher ToM scores. More recently, Peters et al. (2016) found that a sample of individuals who experienced non-clinical auditory hallucinations reported better mindfulness skills compared to individuals diagnosed with psychosis. These studies indicate that robust mentalizing may represent a factor that protects at-risk individuals from the psychopathological effects of emerging psychosis during its preclinical stages. Despite these findings, the nature of associations linking mentalizing and emerging risk for psychosis during the premorbid and prodromal stages of the illness remains incompletely understood.

The second part of the chapter will explore the hypothesis that disruptions in ToM and reality/self-monitoring, which are commonly identified in people suffering from clinical psychosis, can also be observed during the preclinical stages of the illness, both at the state and trait levels of its expression. Furthermore, it will be shown the expression of schizotypal traits during late childhood and adolescence is linked with early disruptions in these mentalizing

processes. From this basis, it will be argued that the expression of schizotypal traits in adolescence may causally impact on the normative development of mentalizing processes involved in the understanding of the self and others (i.e. ToM and self-monitoring), thus augmenting the risk for the development of clinically-relevant symptoms. To frame the subsequent discussion, the next section begins by briefly outlining the concept of mentalizing in terms of its developmental nature, suggesting that childhood and adolescence represent critical periods both for the acquisition and developmental elaboration of mentalizing capacities, particularly in the context of close interpersonal relationships with others.

1.4.1. Understanding oneself and others: the acquisition of mentalizing within interpersonal relationships with others in childhood and adolescence

The capacity to attribute mental states in oneself and others is increasingly understood as constituting a developmental achievement linked closely to the quality of interpersonal relationships during childhood and adolescence (Fonagy et al., 2002). Results from empirical investigations have supported the view that early attachment relationships exert a developmental impact on the understanding of mental states. For instance, Fonagy, Redfern and Chapman (1997), reported that the level of attachment security, measured by a separation anxiety test (Klagsburn and Bowlby, 1976) was positively correlated with ToM performance among a sample of children aged between 3-6 years. In a more recent study using a prospective design, Steele, Steele and Croft (2008) tested basic facial emotion recognition capacities in a sample of 6 year old children (n = 63) that had previously undergone an assessment of mother-infant attachment (Strange Situation Procedure, Ainsworth et al., 1978) at one year of age. The authors reported that 6-year-old children who were classified as having an insecure attachment to their mother at 1 year of age were significantly less accurate in their emotion recognition judgments compared to children that were classified as securely attached. Interestingly, results

from the study also showed that the understanding of emotions by 6-year old children was unrelated to both children's and parents' verbal skills. Thus overall, the results by Steele, Steele and Croft (2008) suggest that the acquisition of mentalizing skills pertaining to emotion recognition represents, at least early in development, a pre-verbal phenomenon linked primarily to the quality of parent-infant relationships. On this basis, developmental models have sought to explain how subtle embodied and psychological interactions between children and their parents may contribute to foster the acquisition of mentalizing and sustain the development of a coherent sense of embodied and psychological selfhood.

At the most basic level, infantile attachment-seeking behaviours (i.e. proximity seeking, clinging) are typically perceived and reciprocated by attuned caregivers via closely, but not perfectly, contingent physical actions (i.e. holding, soothing, feeding). By sensitively responding to expressions of physical needs, the caregiver facilitates in the infant the emergence of a pre-reflective sense of agency, in that self-produced physical actions are experienced as functions that bring about goal-directed outcomes in the external environment (Csibra & Gergely, 1998). Repeated embodied interactions with caregivers affirm the reality and validity of infants' physiological arousal, thus allowing them to integrate their body-related experiences in order to scaffold a basic sense of themselves as separate human beings (Fonagy & Target, 1997; Fotopoulou & Tsakiris, 2017). Indeed, indirect support for the role of early embodied interactions in the physical regulation of the self can be found in studies suggesting that increased early skin-to-skin mother-infant stimulation is associated with better visual-motor skills, and advanced motor development in low birth weight infants (Weiss, Wilson and Morrison, 2004), as well as better body-temperature maintenance and stress regulation in newly-born infants following Caesarean section (Stevens et al., 2014).

Furthermore, by responding psychologically through mirroring responses (attuned yet slightly distorted reflection through facial expressions and tone of voice) to their children's

physical expressions of complex and overwhelming emotional states (particularly those of distress), caregivers provide them with an implicit sense that these can be communicated, recognised and managed. This highly intuitive dyadic interaction provides children with an experience of "*containment*" (Bion, 1962) that modulates their affective states and facilitates in them the capacity to mentally represent, give meaning to and take ownership of self-experience. Repeated experiences of mirroring complex emotions by caregivers enables children to find meaning in their affective states and internalize them as part of their self-regulatory repertoire. In this way, young children transition from split modes of experiencing their inner states, as either isomorphic to external reality (*psychic equivalence*) or completely dissociated from it (*pretend mode*), to a more mentalizing or *reflective mode* where mental states are experienced as inherently meaningful and subjective in nature (Fonagy & Target, 1996). Thus, with the help of ordinary attuned parenting, children learn to use their minds (i.e. their mentalizing capacity) to recognise their inner states and provide for themselves the caregiving function originally afforded by their parental figures. Indeed, the quality of caregivers' mirroring during the first six months has been empirically shown to influence children's capacity to regulate affect (Gergely, 2004). Furthermore, as mentioned in the beginning of this section, a number of studies have shown that secure attachment in infancy and early childhood is positively associated with the normative development of mentalizing processes such as ToM, emotion recognition and mindfulness (Fonagy et al., 1997; Sharp & Fonagy, 2008). In a similar vein, neuroscientific evidence indicate that as adults we recruit the same neural networks when engaging in parent-referential and self-referential reflective processing (Vanderwal et al., 2008), thus supporting the role of early affect-mirroring in mentalizing (Fonagy & Luyten, 2009).

While early attachment relationships during infancy and childhood are considered critical for the establishment of mentalizing capacities (Fonagy & Target, 1996), increasing

evidence suggest that adolescence represents a key developmental period for the maturation and elaboration of mentalizing processes (Fonagy et al., 2002). At the neural level, data suggest that significant structural and functional brain changes take place in key networks implicated in mentalizing during adolescence (Blakemore, 2008; Choudhury et al., 2006). These changes have been shown to significantly contribute in the cerebral specialization of mentalizing processes (Badoud et al., 2016). Simultaneously, within the interpersonal domain the lessening of young peoples' dependence on family systems and the initiation of new interpersonal relationships (e.g. peer, romantic) lead to novel affective experiences (e.g. sexual arousal, shame, etc.), which further contribute to the elaboration of mentalizing processes as a means to navigate a social world of increasing complexity (Fonagy et al., 2002). This is supported by a growing number of behavioural data suggesting that mentalizing performance continues to improve throughout adolescence and into young adulthood (Choudhury et al., 2006; Dumontheil et al., 2010; Vetter et al., 2013). Therefore, the capacity to construe representations of one's own and other peoples' mental states appears to be shaped both through adolescent brain development, as well as through the wealth of novel social learning opportunities that typically arise in adolescence (Fonagy et al., 2002; Poznyak et al., 2019).

Critically however, the rapid biopsychosocial changes that take place during adolescence (e.g. hormonal changes, synaptic pruning, increasing importance of school performance, identity formation and individuation processes), along with the interpersonal and emotional stress-load these typically generate, also constitute it as a critical period for the emergence of psychopathological manifestations (Paus et al., 2008). Indeed, schizotypal trait manifestations commonly emerge in early adolescence (Debbané & Barrantes-Vidal, 2014) and become clinically relevant during the latter stages of adolescent development (Gooding et al., 2005). Importantly, behavioural and neuroimaging studies indicate that the expression of schizotypal personality traits and psychotic-like experiences in young people is associated with

adverse interpersonal outcomes, as well as structural alterations and atypical neural activation in brain areas involved in mentalizing processes pertaining to the understanding of the self and others (Romero-Garcia., 2020; Dahoun et al., 2013; Lagioia et al., 2011). These findings signify that for some adolescents, the development of mentalizing occurs in the context of trait risk for psychosis. Indeed, recent findings suggest that young people who display subclinical psychotic manifestations demonstrate reduced developmental trajectories for common age-related improvements in ToM compared to healthy controls (Davidson et al., 2018). As such, one possibility is that subtle neuro-cognitive and interpersonal aberrations arising in the context of schizotypal trait expression may interfere with the normative development of mentalizing processes during adolescence and impair the capacity for self and other understanding. Disruptions in mentalizing may in turn lead to the exacerbation of schizotypal manifestations in a developmental transactional process that augments psychosis vulnerability across adolescence (Debbané & Barrantes-Vidal, 2014).

Interestingly, empirical evidence has shown that the risk for the secondary development of clinically-relevant symptoms among young people who report non-clinical perceptual aberrations is higher among those who also exhibit worse ToM abilities (Bartels-Velthuis et al., 2011). These findings suggest that the level mentalizing ability, particularly as it pertains to understanding the mental states of others, may represent a factor that interacts with schizotypal traits to moderate the latter's effects on the development of clinically-relevant symptoms. It must be noted however that considering mentalizing as a moderator of the relationship between schizotypal traits and psychosis risk potentially ignores that schizotypal traits and mentalizing are not independent from each other. Indeed, as it will be reviewed in more detail throughout the next sections, a number studies have reported significant associations between the two. On this basis, an empirically unexplored question pertains to whether disruptions in mentalizing processes represent a mediating pathway through which

schizotypal trait manifestations indirectly impact on the development and experience of clinical symptoms. According to this hypothesis, aberrations in the perceptual, cognitive and interpersonal domains that arise in the context of schizotypal personality traits, may contribute to impair the developmental capacity to understand mental states during adolescence, and these impairments may in turn directly relate to the development of clinically-relevant symptoms (Debbané & Barrantes-Vidal, 2014; Salaminios et al., 2021). At present, little is known about the nature of associations linking schizotypal personality traits to mentalizing difficulties during the critical developmental period of adolescence. Furthermore, the significance of mentalizing disfunction on the development of clinical symptoms among young people who exhibit trait risk for psychosis remains incompletely understood.

The next sections will review evidence suggesting that impairments in mentalizing processes pertaining to self and other understanding, typically identified in people diagnosed with clinical psychosis and help-seeking individuals who meet CHR criteria, are also associated with the expression of schizotypal traits in non-clinical samples. Particular emphasis will be given on the small number of studies suggesting that mentalizing difficulties in the context of high psychometric schizotypy are already observable during childhood and adolescence. It will be argued that research on schizotypal trait expression during adolescence can elucidate our understanding regarding the early course and significance of mentalizing dysfunction during the earliest stages of emerging psychosis with important implications for early prevention treatments.

1.5. Understanding the mental states of others: Theory of Mind (ToM) in the psychosis continuum

Theory of mind (ToM) refers to the capacity to make inferences about other people's mental states in order to understand and predict their behaviours (Premack and Woodruff, 1978). ToM is a multifaceted psychological process involving both cognitive and affective components.

Affective ToM relates, at least in part, to the concept of empathy, which involves both relating to and recognising the emotional states of others (Shamay-Tsoory et al. 2007). Cognitive ToM on the other hand captures the capacity to form an understanding of other people's thoughts and intentions on the basis of their overt behaviours (Fonagy & Luyten, 2009). The conceptual distinction between these two ToM dimensions is supported by neuroimaging data showing differential patterns of prefrontal cortex activation for cognitive and affective mental state inferences (Shamay-Tsoory et al., 2007). More specifically, studies suggest that inferences about other people's thoughts and intentions are underpinned by cortical midline prefrontal structures, while inferences about their emotional states are based on lateralised regions, such as the ventromedial prefrontal cortex (Shamay-Tsoory et al., 2005; Acosta et al., 2019). Most experimental tasks designed to assess basic mental state understanding commonly measure performance in singular components of ToM. For instance false belief (e.g. Strange Stories task; Happé, 1994) and perspective-taking tasks (e.g. Director task, Dumotheil et al., 2010) mainly tap into cognitive aspects of ToM. Conversely, emotion recognition tasks (e.g. Reading the Mind in the Eyes Test, Baron-Cohen et al., 2001) also require an empathic appreciation of emotional states in others (Shamay-Tsoory et al. 2007). Table 1.3. presents a list of tasks typically used to test ToM performance.

Task name	Content	Stimuli presentation	ToM domains tested
Strange Stories Task (Happé, 1994)	24 short vignettes presented verbally, each accompanied by a picture. Participants are asked to explain why a character says something that is not literally true	Verbal/pictorial	False-belief understanding
Director Task (Dumotheil et al., 2010)	Participants are presented with virtual images of a bookshelf and are asked to move objects around by interpreting the instructions given by a partner who can only view some of the objects	Pictorial	Visual perspective-taking
Sally-Anne Test (Baron-Cohen et al., 1985)	Participants are presented with picture sequences of one character placing an object and another character (unknowingly to the first character) moving the object. Participants are asked where the first character will look for the object	Pictorial	Faux pas
Reading the Mind in the Eyes Test (Baron-Cohen et al., 2001)	Participants are presented with photographs of male and female eyes depicting emotional states and are asked to select the emotional state that best describes the eyes	Pictorial	Emotion recognition
Movie for the Assessment of Social Cognition (Dziobek et al., 2006)	Participants are asked to watch a 15-minute video of four characters interacting and are asked to infer their mental states (thoughts, emotions, intentions) at different points during the interaction	Video	False belief, faux pas, metaphor, sarcasm, irony, emotion recognition

Table 1.3. Summary description of selected tasks commonly used to assess ToM

1.5.1. ToM in clinical psychosis

Meta-analytic data from a large number of studies have shown that patients with established schizophrenia exhibit worse performance in ToM tasks compared to non-clinical controls (Brüne, 2005) and these appear to be independent of general intelligence, age and gender (Pickup and Frith, 2001; Harrington et al., 2005; Sprong et al., 2007). Furthermore, evidence suggest that impaired ToM is linked to the severity of psychotic symptomatology in those

suffering with the illness (Frith, 2004). Interestingly however, studies exploring the associations between specific symptom sub-groups and ToM have generated inconsistent results. For instance, a number of studies have found that poor performance in ToM tasks is primarily associated with negative symptomatology (Corcoran and Frith, 2003; Corcoran et al., 1995; Weijers et al., 2018), while others have reported associations with positive (Sarfati et al., 1997; Frith and Corcoran, 1996; Versmissen et al., 2008) and disorganization symptoms (Sarfati et al., 1999; Mazza et al., 2001). These inconsistent findings may reflect the outcomes of limitations in the methodologies used to assess the associations between ToM and specific symptom subgroupings. For instance, most studies that reported associations between ToM and negative symptoms did not control for the effects of illness chronicity. This is important as illness chronicity has previously been linked with the severity of negative symptoms and ToM performance (Langdon et al., 2001). As such, the relationship between ToM and negative symptoms identified in some studies may have emerged due to the length of illness in the studied sample (Harrington et al., 2005). Most importantly perhaps, very few studies that have examined the relations between ToM and specific symptom clusters have accounted for the possible contribution of other symptoms to their findings (Harrington et al., 2005). Indeed the hierarchical subgrouping method used by certain studies would place the negative symptom subgroup at the highest category, which meant that patients in this subgroup could also report positive symptoms, but those in the positive symptom subgroup would not report negative symptoms (Sprong et al., 2007). Despite these methodological limitations, most evidence suggest that impaired ToM relates to the phenotypic manifestations of the illness.

It has previously been proposed that ToM difficulties in psychosis may be explained by multiple underlying mechanisms and these may also characterise the symptomatic expression of the illness (Frith, 2004). According to Frith (2004), patients who exhibit predominantly negative symptoms and disorganized thinking, experience a deficit in the representational

abilities necessary to attribute mental states, thus display a *hypomentalizing* pattern characterised by reduced use of inner mental states to understand others' behaviours. Conversely, patients who primarily exhibit positive symptoms, such as paranoid delusions or delusions of reference, may possess the basic representational abilities needed to understand mental states. However, they tend to apply these in a biased way by overattributing intentional mental states or self-referencial meaning to others' behaviours in the absence of observable data to support their inferences (Frith, 2004). This latter patern has been referred to in the literature as "hypermentalizing" (Sharp et al., 2011) or "hyper-ToM" (Clemmensen et al., 2014) and has also been observed in adolescents exhibiting borderline personality features (Sharp et al., 2011; Sharp et al., 2013).

Only a small number of studies to date that have utilized task-based measures specfically geared to distinguish between hypermentalizing and hypomentalizing errors in ToM. These have found differential associations with positive and negative symptom dimensions. For example, Montag et al (2011) compared ToM performance between patients suffering from schizophrenia (n = 80) and a group of healthy controls (n = 80) using the Movie for the Assessment of Social Cognition (MASC, Dziobek et al., 2006), a computerised video-based task that allows for the study individual differences in ToM as they unfold withing complex social interactions by assessing both hypomentalizing and hypermentalizing errors. In line with the hypotheses, results showed that the clinical group exhibited significantly more hypomentalizing and hypermentalizing errors compared to healthy controls. Furthermore, within the clinical group, negative symptoms were associated with an increased number of hypomentalizing errors, while positive symptoms were associated with an increased rate of hypermentalizing ToM errors. The later finding confirms that the heterogenous symptomatic manifestations observed in schizophrenia may be linked to different types of ToM dysfunction (i.e. hypomentalizing vs. hypermentalizing). Most interestingly perhaps, this study offers

empirical support to the proposed relation between hypermentalizing attributions and the positive symptoms of schizophrenia-spectrum illnesses.

1.5.2. ToM in CHR samples

Empirical research has further shown that ToM impairments can be identified among individuals experiencing their first episode of psychosis (FEP) and these are comparable to those seen in patients with chronic psychosis (with effect size Cohen d = 1.0) (Bora and Pantelis, 2013). These findings suggest that ToM difficulties are already present during the initial expression of clinical psychosis, thus do not constitute the outcomes of illness progression and chronicity, or reflect the effects of long-term pharmacological treatment. Indeed, data suggest that CHR individuals who exhibit state manifestations of psychosis risk (i.e. attenuated and intermittent psychotic symptoms), prior to the development of full-blown symptoms, also display worse performance in ToM tasks compared to non-clinical controls. For example, Chung et al. (2008) reported significantly worse ToM performance in a CHR sample of young adults compared to age- and IQ-matched healthy controls, with moderate effect sizes (0.64 - 0.68).

A critical question raised by these findings pertains to whether ToM dysfunction during the prodromal risk phase of psychosis expression represents a predictive risk factor for the development of clinical illness (Bora and Pantelis, 2013; Thompson et al., 2012; Thompson et al., 2011). Yet only a limited amount of research to date has examined the predictive effects of ToM dysfunction on the transition from clinical high risk states to the psychiatric forms of the illness, yielding mixed results. For instance, Kim et al. (2011) found that CHR individuals who transitioned to a first episode of psychosis over a 5-year period showed worse baseline scores in tasks assessing multiple domains of basic ToM abilities compared to individuals who did not convert to clinical illness. This is in line with more recent findings showing that baseline

mentalizing difficulties, as assessed by a narrative interviewer-rated measure (Reflective Functioning Scale, Fonagy et al., 1998), were prospectively related to an increased likelihood of developing clinical psychosis (OR = 0.43) in CHR adults (Boldrini et al., 2020). Contrary to these findings however, longitudinal data from a large cohort of CHR young adults (n = 764) indicate that differences in ToM alone may not be sufficient to predict conversion from the prodromal stage of psychosis-risk to clinical psychosis (Pisculic et al., 2016). Although, in accordance to previous findings, the CHR group in the particular study displayed significantly worse performance in ToM compared to healthy controls, within the CHR group those who converted to a first episode of psychosis did not display differences in ToM compared to non-converters (Piskulic et al., 2016). Nonetheless, even if ToM differences alone were not sufficient to predict conversion from the prodromal stage of psychosis-risk to clinical psychosis, improvements in ToM over time were shown to be significantly more pronounced among CHR individuals who did not transition to clinical psychosis compared to those who did (Piskulic et al., 2016).

Prospective longitudinal studies are important to further our understanding regarding the early social cognitive patterns that precede the full-blown onset of psychotic illnesses (Debbané, Salaminios et al., 2016). Although current research indicates that ToM impairments in CHR individuals may not constitute consistent "predictors" of transition to clinical psychosis, they do appear to represent non-specific vulnerability factors for the development of the illness among those who are at increased clinical risk. Indeed, it has recently proposed that mentalizing dysfunction may not be causally associated with clinical psychosis, but rather signify the breakdown of resilience protecting against transition to clinical illness among those who are at increased risk (Debbané, Salaminios et al., 2016). Further evidence in support of the link between ToM and emerging risk for psychosis come from studies in non-clinical adults, children and young people who report schizotypal personality traits.

1.5.3. The relationship between ToM and schizotypal personality traits in adults

A relatively small amount of cross-sectional research has examined the links between ToM and schizotypal trait dimensions in adult samples. Langdon and Coltheart (1999), found selective ToM impairments in a picture sequencing task (assessing false-belief and social perception understanding) among non-clinical individuals reporting high scores on measures of positive schizotypy. These were found to be independent of executive planning or inhibitory control deficits. Furthermore, in the only correlational study in an adult sample to date, Pickup (2006) assessed the relationship between schizotypy dimensions and ToM performance as measured by the Strange Stories Task (Happé, 1994) in a sample of non-clinical adults. High scores on the cognitive-perceptual dimension of schizotypy in this sample were associated with reduced ToM scores.

In line with studies in clinical samples, a hypermentalizing ToM pattern has also been absorved among non-clinical individuals reporting high scores on measures of positive schizotypal manifestations and delusion proneness. Fyfe et al (2008) presented a group of healthy adults with a computerised task in which participants were asked to view three animated triangles move around the screen under three different conditions: random (pusposeless movements); goal directed (i.e. fighting or dancing); and a ToM condition likely to evoke mental state attributions (i.e. one triangle trying to persuade another one to set it free). In each condition of the task, participants were asked to describe what they perceived to be happening in the animation. Hypermentalizing was calculated as the sum of random condition items (i.e. pursposless movements), in which a reference was made in the participants' description about thoughts, feelings, beliefs and intentions of either the triangles, or about a purposeful interaction between the triangles. Results of the study showed that individuals scoring high on measures of delusion proneness and positive schizotypal trait expression, exhibited significantly more hypermentalizing errors compared to those with low scores on

these measures (Fyfe et al., 2008). These findings suggest that hypermentalizing ToM missatributions are not only found in samples of people suffering with clinical psychosis, but may also relate to trait signs of psychosis risk. However, it must be noted that contrary to Montag et al (2011) who assessed ToM processes in clinical psychosis as these unfold within complex, real-life social interactions, the triangles ToM task used by Fyfe et al., (2008) required partipants to undertake mental state attributions in the context of highly artificial and basic interactions, a factor that may have impacted the data.

1.5.4. The relationship between ToM and schizotypal personality traits during childhood and adolescence

The relationship between hypermentalizing ToM inferences and early trait signs of psychosis risk has also been observed during early development, among non-clinical childen from the community. Clemmensen et al. (2014) examined ToM and its association with non-clinical psychotic-like experiences in a sample of 1630 children from the general population (11-12 years of age). It was hypothesised that a lifetime history of non-clinical psychotic-like experiences (i.e. hallucination- and delusion-like phenomena) in this sample would be associated with a selective hypermentalizing ToM pattern. Hypermentalizing in this study was assessed as a categorical variable measured on the basis of scoring among the lowest 50% of the sample with regard to overall ToM abilities, while concurrently scoring on the top 50% of the sample with regard to giving mentalizing answers refering to beliefs, desires and other inner states. In terms of overall ToM abilities, results showed that children with low ToM scores (i.e. scoring below the median value) were more likely to have experienced psychotic-like experiences compared to children with ToM scores above the median. Most importantly however, children who displayed a hypermentalizing pattern were significantly more likely to have reported psychotic-like experiences than non-hypermentalizing children (OR = 2) and

this effect was particularly pronounced for experiences of paranoid/persecutory ideation (Clemmensen et al., 2014).

In a more recent study in the same sample, Clemmensen et al. (2016) further examined the specificity of the association between hypermentalizing and a lifetime history of non-clinical psychotic-like experiences compared to the later's relation with other known factors linked to psychosis risk, such as family illness, concurent psychiatric diagnoses, bullying, gender and changes in socio-economic status. Results of the analyses showed that hypermentalizing along with concurent psychiatric diagnoses, involvement in bullying and low family income were all associated with psychotic-like experiences. Critically though, hypermentalizing was the sole factor that was independently and significantly associated with psychotic-like experiences in the absence of concurrent psychiatric diagnoses. Interestingly, although most factors associated with psychotic-like experiences in this study were also linked with non-psychotic general psychopathology (e.g. bullying), hypermentalizing alone was specifically linked with the expression of non-clinical psychotic-like manifestations (Clemmensen et al., 2016).

Surprisingly, only one study to date has directly tested the relationship between ToM performance (false belief understanding assessed using the Strange Stories task) and schizotypal trait dimensions in a non-clinical adolescent sample (13-16 years of age, mean age = 14.51) (Barragan et al., 2011). In line with data from adult samples (Pickup, 2006) this study also found an inverse association between ToM performance and scores on the cognitive-perceptual dimension of a schizotypy scale that included trait features of perceptual aberrations, magical thinking and hallucinatory experiences (Barragan et al., 2011). Furthermore, in accordance to adult schizotypy studies, but contrary to studies in individuals suffering with clinical psychosis, no associations were found between ToM performance and the negative or disorganized schizotypy dimensions.

Most interestingly perhaps, studies suggest that the effects of schizotypal traits on the development of clinically-relevant symptoms may depend upon the level of ToM dysfunction. For instance, Bartels-Velthuis et al. (2011), reported that in sample of children (aged 7-8 years) and young adolescents (aged 12-13 years) who experienced non-clinical hallucinations within the previous 5 years, the risk for the development of secondary delusional ideation was higher for those with lower ToM scores. The task used in this study was the ToM Storybook Frank (Blijd-Hoogewys et al., 2008), in which participants are presented with pictures of a story or listen to the story read out loud and are tested on the ability to understand first and second-order false beliefs, white lies, deception, double bluffs and faux pas.

Overall, data from the small amount of studies in non-clinical children and adolescents from the general population are suggestive of a link, quite early in development, between ToM and the first trait signs of emerging psychotic pathogenesis. Importantly however, Nevertheless, ToM tasks only capture one aspect of the broader multidimensional concept of mentalizing. Indeed, increasing clinical accounts and empirical evidence suggest that difficulties in multiple dimensions of mentalizing, which extend beyond the "other-oriented" focus of ToM, may relate to psychotic phenomena across the clinical and subclinical levels (Debbané, Salaminios et al., 2016; Boldrini et al., 2020). For instance, it has been shown that disturbances in recognising and monitoring one's own mental events (i.e. "self-oriented" mentalizing) may critically underpin the reality distortions characterising psychotic phenomena (Sass and Parnas, 2003; Debbané, Salaminios et al., 2016). This is important as recent evidence suggest that disruptions in self-referential processes, observed both at the phenomenological and neurocognitive levels, can discriminate between schizophrenia and non-schizophrenia spectrum disorders in CHR samples (Spark et al., 2021). Furthermore, it has been proposed that disruptions in self-reflective processes, particularly during the premorbid period of psychosis expression in childhood and adolescence, may contribute in the aberrant

explanations of social experiences (i.e. ToM dysfunctions) commonly identified in people who exhibit clinical state manifestations or non-clinical psychotic-like experiences (Brent and Fonagy, 2014; Debbané, Salaminios et al., 2016). Within this context, an interesting hypothesis is that higher-order mentalizing processes involving the attribution of inner states in others, such as ToM, may be underpinned by self-referential abilities pertaining to the monitoring of the self.

Indeed, emerging data from functional neuroimaging studies in non-clinical subjects suggest that awareness of inner states in the self and others involves a shared neural network ("the neural circuitry of the self"; Frith and Frith, 2003), which includes the medial prefrontal cortex, the superior temporal sulcus and the temporal poles bilaterally (Murphy et al., 2010; Brent et al., 2009). Interestingly, hypoactivation patterns in this neural system have been observed during both self- and other-oriented mentalizing evaluations in patients suffering with clinical psychosis, as well as in non-clinical adolescents who report schizotypal traits and hallucination-like experiences (Murphy et al., 2010). At the behavioural level, data from a clinical psychosis sample have also shown a strong correlation between impairments in self-referential processing and ToM performance (Fisher et al. 2008). Thus, neuroimaging and behavioural data both suggest that the ToM difficulties commonly identified among clinical psychosis samples and in those who report high psychometric schizotypy may be indicative of a more fundamental disruption in the monitoring of self-generated mental events (Murphy et al., 2010).

The next sections will review studies suggesting that disfunctions self-reflective abilities, namely self- and reality-monitoring (i.e. the capacity to identify and monitor one's self-generated mental events) relate to psychotic phenomena in people suffering from the illness, but also in non-clinical adult and adolescent samples.

1.6. Understanding oneself: self- and reality-monitoring in the psychosis continuum

According to contemporary conceptualizations, clinical manifestations of psychotic illnesses, particularly those that entail cognitive-perceptual aberrations, are associated with confusions in the identification of self-generated mental events. For instance, auditory hallucinations have been linked with difficulties in monitoring the origin one's own inner speech (Allen et al., 2007), while delusions of control have been attributed to difficulties in monitoring the initiation and sensory consequences of self-produced actions (Blakemore et al., 2003). A key metacognitive process associated with the monitoring of self-generated mental events is the ability to discriminate between different sources of memory information, also known as *source monitoring* (Johnson et al., 1993).

Source-monitoring processes are typically divided into three types: (1) *reality-monitoring* involves the capacity to distinguish between self-generated and externally-perceived information (e.g. by another agent); (2) external source-monitoring refers to the ability to distinguish between two external sources of information (e.g. between words spoken by a male or female voice); and (3) *self-monitoring* involves the capacity to discriminate between two internal sources of information (what one imagined doing or saying from what one did or said). To study these processes, Johnson et al., (1993) developed the *source-monitoring framework*, a conceptual approach seeking to explain the parameters that promote accurate source-monitoring discriminations, as well as factors that underlie source-monitoring confusions. According to this model, judgments about the source of memory information are based on evaluations pertaining to the characteristics of the memory events themselves. These include the contextual (spatial/temporal) and sensory characteristics (visual/auditory) available, as well as the record of cognitive operations engaged during the original encoding of an event, or evoked during memory retrieval. For instance, studies suggest that memories for externally-perceived events contain more sensory information compared to memories for

imagined events and the capacity to discriminate between the two increases with the amount and clarity of such detail contained in the former (Dewhurst, 1999; Johnson, 1988). Conversely, because imagination is less automatic than perception, memories for imagined events contain more records about the cognitive operations that led to their creation compared to those for real events (Sussman, 2001). Indeed, increased cognitive effort during memory encoding for imagined events has been shown to support accurate source-monitoring discriminations (Finke et al., 1988).

Research suggests that developmental changes in certain domains of source-monitoring abilities occur between childhood and young adulthood (Markham, 1991). For example, Foley et al (1983) found that 6-year-old children showed a higher rate of self-monitoring memory errors compared to 17-year-olds when asked to discriminate between what they had said and what they had imagined themselves saying. Similarly, Markham (1991) compared children of 6, 9 and 12 years of age on an action self-monitoring task, with the results showing that 6-year-old children experienced more confusions in discriminating between self-generated performed and imagined actions compared to older children. This line of evidence suggests that the capacity to discriminate in memory between what one said or did and what one imagined saying or doing, undergoes further elaboration from early childhood to young adulthood. Interestingly, little is known at present about the developmental trajectories of self-monitoring processes during adolescence.

In contrast to self-monitoring, developmental research on reality-monitoring suggests that the capacity to differentiate in memory between self-generated and other-produced actions is established in early childhood. For instance, Folley et al (1983) found that 6-year-old children performed as good as 17-year-olds in discriminating between memories of what they had said from memories of what they had heard another person saying. Therefore, although neuroimaging data suggest shared neural activation patterns during reality- and self-monitoring

evaluations (Lagioia et al, 2011), behavioural findings suggest that the two processes may undergo different developmental trajectories.

Regarding impairments in self- and reality-monitoring, the source-monitoring framework postulates that because contextual characteristics and cognitive operations are used as cues that determine the origin of memory events, source confusions are likely to occur under conditions that alter the quantity and quality of such information during encoding (Johnson et al., 1993). For example, studies indicate that conditions that increase the sensory detail encoded for imagined events may lead to their misattribution as externally-perceived and real (Johnson, 1988). Similarly, increases in the cognitive elaboration required during the encoding of externally-perceived events may lead to their misattribution as self-generated and imagined (Johnson, 2006). Although source memory is an imperfect process and transient source confusions are relatively common, impairments in source-monitoring, particularly with regards to self- and reality-monitoring processes have been consistently linked with psychotic phenomena, both at the clinical and subclinical levels of expression.

1.6.1. Self-monitoring in clinical psychosis and CHR states

Given that psychotic experiences entail pervasive difficulties in discriminating between imagination and reality, a number of studies have investigated self-monitoring processes in people diagnosed with clinical psychosis. These commonly aim to assess participant' ability to distinguish between two internal sources of information (what one did from what one thought or imagined doing).

In one of the most common experimental paradigms designed to test self-monitoring performance, participants are asked to read, either aloud (external condition) or silently (internal condition) a series of word-items. In the test phase, after an interval delay, participants are presented with the word-items from the first phase of the task, along with a series of new distractor word-items and are asked to indicate if each item had been previously presented

(recognition test), as well as attribute recognized items to a reading condition, silent or aloud (self-monitoring test). This type of task yields two types of self-monitoring errors: externalizing errors (attributing silently-read items as overtly-read), and internalizing errors (attributing overtly-read items as silently-read).

Using this experimental paradigm, Keefe et al. (1999), reported that a group of patients suffering with schizophrenia (n = 28) exhibited significantly more externalizing self-monitoring errors, in terms of confusing in memory word items they had read silently as overly read, compared to a group of healthy controls (n = 19). Franck et al. (2000) used the same experimental paradigm to compare self-monitoring performance between a group of patients suffering from auditory hallucinations (n = 17) and a group of healthy controls. In line with the findings reported by Keefe et al. (1999), the patient group exhibited significantly more externalizing errors compared to the control group. In addition, the patient group also exhibited the tendency to misattribute the newly presented distractor items as overtly read. According to Franck et al. (2000) these findings suggest that the tendency to confuse inner speech as real may cognitively underpin the experience of auditory hallucinations. In a similar study by Henquet et al (2005), a group of patients diagnosed with schizophrenia (n = 15) and a group of non-psychiatric controls (n = 15) were compared in terms of performance on a self-monitoring task in which verbalized thoughts had to be discriminated from silent thoughts. The task consisted of twenty four trials in which participants were presented with simple questions on a computer screen for which they were required to provide an answer. Half of the trials involved the presentation of a single question, while on the other half were presented in pairs. In the paired-questions condition, participants were required to attend to and answer both questions, but were asked to only verbalize their answer for one. As such, certain questions received verbal answers (external condition) while the rest were answered covertly in the absence of verbalization (internal condition). Relative to the control group, patients with schizophrenia

exhibited significantly more externalizing self-monitoring errors by misattributing in memory answers that were given silently as overtly verbalized.

Overall, the studies reviewed above suggest that patients suffering with clinical psychosis display a tendency towards misattributing in memory silent or imagined speech as overtly verbalized. Furthermore, evidence suggest that the tendency to confuse between silent and overt speech may correlate with the symptomatic manifestations of the illness. Docherty (2012) examined self-monitoring performance in terms of discriminating between silent and overt speech, as well as linear associations with symptom dimensions in a group of patients diagnosed with schizophrenia and schizoaffective disorders (n = 110). In the experimental paradigm used a series of 16 incomplete sentences were presented sequentially to participants. The last word in each sentence was left blank and participants were asked to generate the missing word. For half the sentences participants were asked to say the word out loud, while for the other half they were required to only think the word silently. In accordance with other studies, results showed that compared to healthy controls, the patient group made more self-monitoring errors in memory (sum of externalizing and internalizing misattributions) when asked to discriminate between words they had read aloud and words they read silently, (n = 23). Furthermore, within the patient group, the rate of self-monitoring errors was positively and significantly correlated with the level of self-reported auditory hallucinations and conceptual disorganization.

Importantly, the only self-monitoring study to date undertaken in a sample of young adults at clinical high-risk for the illness (CHR) has shown that similar to patients suffering with psychosis, the tendency to misattribute the source of self-generated mental events to external agents can also be observed among people who display state manifestations of psychosis risk. More specifically, Gawęda et al. (2017) compared self-monitoring performance between a group of prodromal young adults identified on the basis of UHR criteria (n = 36,

M_{age} = 19.17), a group of first-episode psychosis sufferers (FEP, n = 25; M_{age} = 20.36) and a group of healthy controls (n = 33; M_{age} = 20.27) using an action-monitoring task. Participants were presented with either verbal instructions or nonverbal pictograms cuing them to either physically perform or imagine themselves performing certain physical actions (i.e. "raise the glass"). In the test phase participants were asked, after an interval delay, to indicate whether each previously presented action-cue had appeared verbally, or as a pictogram and whether it was physically performed or imagined. Furthermore, participants were asked to grade their confidence on the self-monitoring judgment they made (categorical scale – unsure vs sure). Results of the analyses showed that both the UHR and FEP groups misattributed previously imagined actions as physically performed (i.e. externalizing errors) significantly more than healthy controls and this applied to both verbally- and pictorially-presented action-cues. Furthermore, both UHR and FEP groups evaluated their false responses with significantly higher confidence compared to the healthy control group. These findings suggest that the externalizing self-monitoring tendency observed among individuals suffering with schizophrenia can also be identified among individuals who exhibit state manifestations of psychosis risk.

1.6.2. Self-monitoring and its links to schizotypal trait expression in non-clinical adults and high-risk adolescents

Similarly to research in people at clinical high risk for psychosis, self-monitoring studies in schizotypy have primarily used action- rather that verbal-monitoring tasks, in which participants have to distinguish between self-generated imagined and performed physical actions. (Humpston et al., 2017; Peters et al., 2007). Peters et al. (2007) examined the relationship between action self-monitoring (distinguishing between performed physical actions versus imagined actions) and positive schizotypy in a sample of non-clinical

undergraduate students (M_{age} = 21.41). Participants were asked to either physically perform an action-statement (i.e. "open the door"), or imagine themselves performing the action. In the test phase participants were given a list of all the action-statements presented in the first part of the task and were asked to recall their source (performed vs. imagined). Results showed that participants reporting high scores on the cognitive-perceptual dimension of schizotypy displayed significantly more externalizing misattributions for actions (confusing imagined actions as overtly enacted), compared to those with low scores.

More recently, Humpston et al. (2017) examined the association of action self-monitoring and schizotypal dimensions in a sample of non-clinical adults (n = 102; M_{age} = 22.30). The authors reported that scores on the cognitive-perceptual dimension of schizotypy were positively associated with the total rate of action self-monitoring errors.

In the only self-monitoring study conducted in a sample of adolescents, Debbané et al (2010), examined self-monitoring for speech in 20 young people (mean age = 13.97) prone to express positive schizotypal manifestations due to 22q11.2 deletion syndrome, a neurogenetic condition conferring high-risk for schizophrenia. Their performance on the self-monitoring task was compared with a group of IQ and age-matched controls (n = 19) and a group of age-matched typically-developing controls (n = 19). Participants were presented with both familiar words (low cognitive effort) and non-words (high cognitive effort) and were asked to either read them aloud or read them silently. Data showed that in comparison to both control groups, adolescents with 22q11.2.DS exhibited a greater rate of externalizing self-monitoring speech misattributions. Furthermore, within the 22q11.2DS group, the tendency to misattribute silently read word items as overtly read, was particularly pronounced in the high-cognitive effort condition (Debbané et al., 2010). These findings indicate that high risk adolescents exhibit the same type of self-monitoring confusion to clinical adult samples. Furthermore, they suggest that young people who are at increased risk for psychosis may fail to utilize cognitive

effort as a cue to monitor the source of self-generated covert/imagined mental events. Although these findings are in line with previous research in non-clinical and CHR adult populations, little is known at present about the associations between self-monitoring processes and schizotypal trait expression in typically developing adolescents.

Interestingly, recent findings suggest that self-monitoring confusions can also be identified in adolescents who exhibit non-psychotic psychopathological manifestations. For instance, Morosan et al. (2018), found increased rates of externalizing self-monitoring misattributions in young people with conduct problems (Morosan et al., 2018). Thus, it remains unclear whether self-monitoring difficulties are transdiagnostic and are also associated with other personality traits (i.e. impulsivity) commonly implicated in the development of non-psychotic symptomatic manifestations.

1.6.3. Reality-monitoring in clinical psychosis and CHR states

A number of conceptual and clinical accounts have postulated that the cognitive-perceptual aberrations observed in psychotic disorders are linked with impairments in *reality-monitoring* – that is the capacity to distinguish between self-generated and perceptually-derived (e.g. generated by another agent) information. For instance, auditory hallucinations have been conceptualised as stemming from misattribution of one's own inner speech to a source external to the self (Allen et al., 2007), while delusions of control have been attributed to difficulties in monitoring the initiation of self-produced actions, whose control is instead attributed to external agents (Blakemore et al., 2003).

In the most commonly used reality-monitoring experimental paradigm, the experimenter verbally presents a series of word-items to participants, who are then asked to respond by generating a new word-item of their own (i.e. the name a fruit presented by the experimenter and the participant asked to name another fruit). In the test phase, following an

interval delay, participants are presented with both the experimenter-produced and participant-generated word-items from the first part of the task, and are asked to attribute the source of the recognised word (i.e. experimenter vs. self). Similarly to self-monitoring paradigms, this type of reality-monitoring task yields two types of errors: externalizing errors (attributing self-generated items as experimenter-produced), and internalizing errors (attributing experimenter-produced items as self-generated).

A number of reality-monitoring studies using this experimental paradigm have shown that patients diagnosed with psychotic disorders display significantly more errors towards misattributing the source of self-generated word items to the experimenter compared to healthy controls (Bentall et al., 1991; Brébion et al., 2000; Johns et al., 2010; Johns et al., 2006; Vinogradov et al., 2008; Morrison and Haddock, 1997). Furthermore, this pattern of externalizing reality-monitoring misattributions is shown to be particularly pronounced among patients reporting hallucinations (Bentall et al., 1991; Brébion et al., 2000) and for word-items that are emotionally charged (e.g. positive and negative words) (Morrison and Haddock, 1997)

Another type of experimental paradigm commonly used to test reality-monitoring performance, involves the experimental simulation of participant's actions (i.e. movements, speech). In this type of experimental design, subjects are required to identify their own verbal or motor actions while the sensory feedback of these is experimentally manipulated and distorted. Therefore, contrary to task-based measures described in the studies above, in which participants are asked to distinguish between self- and other-generated events in memory (i.e. after an interval delay), simulation-based reality-monitoring tasks assess subjects' capacity to undertake online (i.e. in-the-moment) reality-monitoring attributions.

Johns et al (2001) used an action-simulation acoustic task to compare reality-monitoring performance between a group of patients experiencing hallucinations and delusions (hallucination group), a group of patients experiencing delusions but not

hallucinations (non-hallucination group) and a group of non-clinical controls. In this task, participants were asked to read aloud a series of words while receiving back through a set of headphones the moment-to-moment acoustic feedback of their own voice under the following conditions: (1) acoustic feedback of their undistorted voice; (2) acoustic feedback with acoustic distortion of their own voice; (3) acoustic feedback of an alien voice (someone else's voice heard); and (4) acoustic feedback with distorted alien voice. Upon reading the word and having listened to the auditory feedback, participants were asked to identify its source (self vs other voice). Results showed that patients suffering with hallucinations exhibited a significantly higher rate of externalizing reality-monitoring errors compared to normal controls when attempting to identify the source of their own distorted voice (misattributing their own voice to an external source), but made fewer errors than controls when identifying the source of an alien voice. Furthermore, the hallucination group exhibited proportionately more externalizing errors for derogatory words, suggesting that patients suffering with hallucinations are more prone to attribute negative material to external sources (Johns et al, 2001). An externalising pattern was also identified, although less prominent, in the group of patients suffering with delusions. On the basis of these findings and in line to studies that involve reality-monitoring judgments in memory, the authors suggested that the experience of hallucinations is linked by an externalization of self-generated speech.

In an action-monitoring experimental designed by Knoblich et al. (2004), participants were asked to draw circles at an average velocity around a static circle displayed on a screen, while a moving dot reproduced the participant's movement on a second screen. After a certain time interval the reproduction of the movement on the second screen was experimentally manipulated with the speed of the movement being increased compared to the actual speed of the participant's movement. Participants were asked to either stop drawing the circle once they noticed the change in velocity on the second screen, or to continue drawing when no change

was detected. According to Knoblich et al (2004), patients suffering with schizophrenia were significantly less sensitive in monitoring the change in velocity between their actual movements and their visual feedback compared to healthy controls. Furthermore, when the patient group was sub-divided on the basis of psychotic symptomatology, individuals suffering with hallucinations and delusions showed decreased sensitivity in detecting mismatches between self-generated actions and their visual consequences, compared to a group of patients without these symptoms. These data indicate that difficulties in monitoring the consequences of self-generated physical actions may underpin core positive symptoms of psychosis (Knoblich et al 2004). In addition these findings, are in line with previous action-monitoring studies also involving patients suffering from psychotic disorders (Franck et al., 2001; Blakemore et al., 2000).

Overall, evidence suggest that samples of individuals who suffer from clinical psychosis consistently display reality-monitoring difficulties characterised by the misattribution of self-generated speech and physical actions to external agents. In addition, these appear to be particularly pronounced in those who exhibit positive symptomatic manifestations. Reality-monitoring research has further shown that similarly to patients suffering with psychosis, individuals at clinical high-risk for the illness (CHR) also display a tendency towards misattributing the source of self-generated items to external agents (Johns et al., 2010). Johns et al (2010) used the auditory distortion source-monitoring task originally utilized with a chronic psychosis sample (Johns et al, 2001 described above), to examine self-monitoring deficits in a group of individuals meeting CHR criteria (mean age = 24.7) and thirty-one age-matched healthy controls. CHR participants were found to be significantly more likely to misattribute their own distorted speech to an alien source compared to healthy-controls, but only under severe distortion conditions. Interestingly, no linear associations were found between reality-monitoring misattributions and specific psychotic manifestations in the

CHR group. According to Johns et al (2010) the symptom data used in this study may have lacked the necessary sensitivity to detect associations between specific symptoms and task performance, while the size of the sample did not allow for analyses based on symptom specific sub-groups. An alternative explanation of this finding offered by the authors is that CHR individuals experiencing more severe symptoms may have been more aware of their difficulties, which may have led them to increase their effort during the source-monitoring task (Johns et al, 2010).

1.6.4. Reality-monitoring and adult schizotypy

The studies reviewed in the previous sections suggest that disruptions in reality-monitoring processes are consistently identified in people suffering with psychotic illnesses. Furthermore, these appear to already be present during the prodromal stage of clinical psychosis expression, among help-seeking young adults who exhibit subclinical state manifestations (i.e. attenuated or intermittent psychotic symptoms). Importantly, reality monitoring misattributions have also been identified in the context of trait risk for psychosis, in community samples and among non-clinical first degree relatives of patients suffering with psychosis. This line of research suggests that reality-monitoring performance relates to the level of self-reported schizotypy, particularly as it pertains to its cognitive-perceptual dimension.

Versmissen et al (2007) used an experimental action-recognition task (Franck et al., 2001; described above) to examine reality-monitoring and its relation to psychosis risk by comparing task-performance between (1) a group of non-clinical individuals with high scores in a measure of schizotypal trait expression (psychometric schizotypy group: n = 40); (2) a genetic high risk group comprised of first-degree relatives of patients suffering with schizophrenia (n = 41); (3) a group of patients with chronic schizophrenia (n = 37); and (4) a non-clinical control group (n = 49). No significant group differences in reality-monitoring were

reported between the clinical psychosis and the high-risk groups in trials that involved the angular manipulation of self-generated stimuli. This finding is in line with the results reported by Franck et al (2001), who only found angular misattribution's among patients suffering from delusions of influence but not among patients with general psychotic symptomatology. However, significant between-group differences were reported in trials that involved the temporal manipulation of self-generated stimuli. Participants in the non-clinical high-risk groups (genetic risk and high psychometric schizotypy groups) showed significant differences to the clinical psychosis and non-clinical control groups for the total rate of action-recognition errors in the specific task. More specifically the non-clinical risk groups exhibited intermediate error rates compared to the psychosis and the non-clinical groups, suggesting a dose-response effect (non-clinical control group: 29.7%, psychometric-risk group: 31.1%, genetic risk group: 31.8%, patients with psychosis group: 37.4%).

Similarly, Hommes et al (2012) used a version of the circle-drawing task (described above) to examine reality-monitoring differences between a group of clinical psychosis patients, their first-degree non-affected siblings and a group of non-clinical controls. In accordance with the data by Versmissen et al (2007), they found that non-affected first-degree relatives of patients with psychosis showed a significantly greater rate of reality-monitoring errors compared to the healthy control group, but a lower rate of errors compared to their affected siblings. Furthermore, it was shown that positive schizotypy in the genetic risk group correlated significantly with the total rate of reality-monitoring misattributions. These findings support data from previous studies suggesting that reality-monitoring difficulties do not simply represent epiphenomena related to the clinical expression of psychosis, but may be aetiologically linked to the development of the illness (Hommes et al, 2012).

Larøi et al (2004) examined the associations between reality-monitoring for speech and hallucination proneness in a sample of non-clinical adults. In the reality-monitoring task

administered, a series of words was presented verbally to participants by the experimenter. Word items varied it terms of emotional valence (10 positive, 10 negative and 10 neutral words). After each word presentation, participants were asked to say the first word that came to their minds. Following an interval delay, subjects were presented with words on a computer screen, consisting of the words originally presented by the experimenter, words generated by the subject, and new distractor words, in random order. Participants were asked to first recall if the word they saw in the screen was old or new. For words that were recognised as old, participants were asked to recall its source (i.e. whether it was generated by oneself or by the experimenter). The sample was subdivided according to their scores on a hallucination proneness scale (the Launay-Slade Hallucination Scale, Launay and Slade, 1981). Participants scoring within the bottom 25th percentile formed the non-hallucination group (n = 25), while those scoring within the upper 25th percentile formed the hallucination group (n = 25). Results showed that the hallucination group made significantly more externalizing misattributions for self-generated word items (i.e. attributed self-generated word items to the experimenter) compared to the non-hallucination group. Importantly, reality-monitoring misattributions for self-generated material differed significantly between the groups for emotionally charged but not for neutral items. Post-hoc analyses within the hallucination group further revealed a significant difference for reality-monitoring misattributions between positive and negative word items (i.e. greater number of misattributions for negative than positive words). The two groups did not differ in the number of internalizing reality-monitoring misattributions (i.e. misattributing experimenter-produced words as self-generated).

Allen et al (2006) examined reality-monitoring for self-generated speech and its association to hallucination proneness and delusional ideation in a sample of healthy adults (n = 57). Participants were first asked to read out a series of adjective words, while their speech was audio-recorded. The voices of male and female researchers reading a series of words were

also audio-recorded, comprising the non-self condition. Participants then listened to the pre-recorded words for which the source (self vs other) and acoustic quality (undistorted vs distorted pitch) differed across trials. In the test phase, participants had to indicate whether the words were spoken in their own or another person's voice. Results showed that participants displayed a propensity towards misattributing their own distorted speech to the experimenter and this was significantly and positively correlated with delusional ideation scores. Contrary to findings from previous studies, only a trend-level positive correlation emerged between externalizing misattributions of the participant's distorted speech and hallucination proneness. Finally, also in contrast to previous studies, the emotional valence of word items did not have an effect on the rate of reality-monitoring misattributions for self-generated material. The authors attributed this on the basis of previous studies suggesting differences in the content of voices experienced by people suffering with clinical psychosis and non-clinical individuals. More specifically, people diagnosed with psychosis often characterise the content of their auditory hallucinations as negative, while non-clinical individuals often report that the voices they hear are non-threatening in nature (Honig et al., 1998). Thus, while people suffering from clinical psychosis may show a tendency towards externalizing their distorted voice when it is negative, non-clinical individuals may not show a similar pattern.

Collignon et al (2005) examined the associations between hallucination proneness and source-monitoring for physical actions in a group of 65 non-clinical individuals (average age 24.2 years). In the source-monitoring task administered, participants were presented with a series of action statements (e.g. "open the door") and were asked to either 1) perform the action, 2) observe the experimenter performing the action, 3) imagine themselves performing the action, 4) imagine the experimenter performing the action, or 5) listen to the experimenter verbally state the action. Following an interval delay, a list of 120 action statements was presented to participants that included the 60 actions that were presented in the study phase

and 60 new actions. Participants were asked to identify if each action was old (previously presented) or new and for actions that were identified as old, to identify their source. Analysis of data showed that hallucination-prone participants exhibited significantly more source-monitoring errors compared to their non-hallucination counterparts, particularly in terms of the two imagined conditions. More specifically, hallucination prone individuals showed a tendency to erroneously attribute actions they had imagined themselves performing to actions they had imagined the experimenter perform.

In a more recent study, Alderson-Day et al. (2019), tested the linear associations between hallucination proneness and reality monitoring performance in a group of non-clinical young adults (n = 76; M_{age} = 20.21 years). Participants were presented with 48 partially completed word pairs (e.g. bacon and e_ _ _) and were asked to either listen to the second word being read out by a male voice, or to complete and read the word pairs out loud. In the test phase the word pairs from the initial phase of the task were presented again and participants were required to decide whether they had heard the word-item or had verbalised the word themselves. In line with previous correlational studies, results of the study did not report linear associations between reality-monitoring performance (i.e. accuracy in recalling the correct source of previously presented items) and hallucinatory experiences. According to Alderson-Day et al. (2019) these findings indicate that the link between reality-monitoring difficulties and non-clinical hallucinations may only be observable among people that are towards the end-points of the psychosis-liability continuum (i.e. people exhibiting high psychometric schizotypy). Indeed, previous studies in non-clinical samples that have found links between reality-monitoring and hallucination proneness have only done so by comparing differences in reality-monitoring between those scoring high on measures of hallucination process and those with low scores on these measures (Larøi et al., 2004; Collignon et al., 2005).

1.6.5. Reality monitoring and schizotypy in adolescent samples

Although reality-monitoring has been extensively studied in clinical samples and non-clinical adults who report high psychometric schizotypy, it remains understudied during the critical developmental period of adolescence. Indeed, only a limited number of studies to date has focused on reality-monitoring processes and their links to schizotypal trait expression during adolescence. However, as it will be discussed below, the reality-monitoring tasks used in these studies focus solely on the capacity to distinguish between self- and non-self-cues "internally" (i.e. distinguish in memory between visualizing oneself perform an action and visualizing another person perform an action), thus do not directly test the capacity to distinguish between self-generated and perceptually-based information.

Debbané et al. (2008), compared source-monitoring performance for actions between a group of young people prone to express positive schizotypal manifestations due to 22q11.2 deletion syndrome, (n = 18; mean age: 14.04), a group of age and IQ-matched controls (n = 17) and a group of age-matched matched healthy controls (n = 17). The authors of the study used an adapted version of the reality-monitoring task used by Collignon et al (2005), described above. Participants in this study were presented with a series of action statements and were asked to either (1) visualize themselves performing the actions; (2) visualize the experimenter performing the action; or (3) mentally repeat the action statement without visualization of the performer. Significant group differences were reported in the action monitoring task, indicating that the 22q11DS group exhibited more source monitoring errors compared to both control groups. Statistical analyses revealed adolescents with 22q11DS tended to misattribute imagine-experimenter actions for actions they had mentally repeated and vice versa. Hence the specific group displayed confusions in discriminating between the conditions of the task that did not involve them visualizing themselves. These findings differ to those of Collignon et al (2005) who found confusions in differentiating between the imagine-self and the imagine-

experimenter conditions of the task in a sample of hallucination-prone adults. A number of reasons may have accounted for the contrasting findings between the two studies. First, the sample examined by Collignon et al (2005) was characterised by hallucination proneness, a factor that was not specifically assessed in the 22q11DS sample. This suggests that externalising source monitoring misattributions may be specifically associated to hallucination proneness in high-risk samples (Debbané et al, 2008). Moreover, given that the frequency of schizotypal manifestations in 22q11DS typically increases with age (Debbané et al, 2008), the inclusion of an adolescent sample may have accounted for the discrepancy in findings between the two studies.

In another study Debbané et al (2009), used the same task to examine the associations between schizotypy and action-monitoring in a sample of adolescents recruited from the general population (n = 163; mean age = 15.3). Results showed that positive schizotypy in this non-clinical youth sample was significantly correlated with the tendency to misattribute imagine-experimenter actions as mentally repeated and vice versa. These results resemble those from the 22q11DS sample that also exhibited significantly more errors of this type compared to age and IQ matched non-clinical control groups (Debbané et al, 2008).

These studies suggest that positive schizotypy in adolescence is associated with a tendency to confuse between visualising an external agent performing an action and mentally preforming the action oneself. As mentioned above however, the specific source-monitoring task only assesses self-other confusions for internal/visualized events, in the absence of an external condition (e.g. the experimenter performing an action), thus does not directly test reality-monitoring. Indeed, the word "reality" alludes to an event's external and perceptually-based nature (Lagioia et al., 2011).

Importantly, this type of source discrimination precludes the processing of sensory information typically engaged when making reality-monitoring judgments for mental events

that have been enacted, thus have adopted an external and perceptually-based quality (Docherty, 2012). For instance, contrary to silent speech, which constitutes a purely internal phenomenon, overt speech also entails the production of motor actions (i.e. vocalizations), along with the perception of the sensory feedback produced by these actions (i.e. listening to the spoken material). Indeed, according to the source-monitoring framework, reality-monitoring misattributions often occur due to alterations in the encoding of sensory information (Johnson et al., 1993). This is in line with the *forward model of motor control* (Frith et al., 2000), which postulates that reality-monitoring confusions result from failures to match the intended or predicted outcomes of self-generated actions to their externally-perceived sensory consequences. This can occur either due to disruptions in generating cognitive predictions about the sensory outcome of actions, or due to impairments in the processing of the sensory feedback related to these actions. In both cases, because of the high level of sensory discrepancy between cognitive prediction and sensory feedback, self-generated events may be experienced as external to the self (Blakemore et al., 2003; Frith et al., 2000). Thus, taken together, the source-monitoring framework and the forward model suggest that the externalizing confusions observed in people who exbibit cognitive-perceptual aberrations may be underpinned by impairments in the processing of sensory and/or cognitive information.

From this standpoint, assessing self-other confusions by using experimental tasks that involve overt speech or actions (also generated by an external agent) would provide a more direct measure of reality-monitoring that could better inform our understanding of its links to schizotypal trait expression during adolescence. In contrast to studies in non-clinical adult samples, no study to date has directly examined the capacity to discriminate self- from non-self-cues for overt speech in typically developing adolescents. Furthermore, previous studies in adult populations have only examined reality-monitoring in the context of cognitive-

perceptual manifestations (i.e. hallucination and delusion-like phenomena). Thus, the relationship between reality-monitoring and the interpersonal or disorganized dimensions of schizotypy remains unclear. This is important given that prospective data indicate that the expression of disorganized and interpersonal dimensions of schizotypy in adolescence is longitudinally linked with the developmental trajectory of clinically-relevant cognitive-perceptual manifestations (Debbané et al., 2013).

1.7. Summary and general discussion of the second part

The second part of the introductory chapter reviewed previous studies that have assessed the capacity to infer mental states in others (ToM), as well as the ability to recognise and monitor one's own self-generated mental events (self- and reality-monitoring), at different stages of the psychosis continuum (i.e. clinical psychosis; CHR; non-clinical schizotypy).

In terms of ToM, disruptions in the capacity to accurately infer mental states in others have been consistently associated with psychotic phenomena at both the clinical and subclinical levels of expression. Although ToM impairments among help-seeking individuals who exhibit state manifestations of psychosis risk do not appear to represent "consistent predictors" of transition to clinical psychosis, they are linked with a non-specific vulnerability for the illness. Furthermore, data from non-clinical populations suggest that ToM difficulties relate to the expression of schizotypal personality traits in adults, but also in samples of children and young people. Surprisingly, contrary to clinical psychosis, studies in non-clinical adult samples suggest that aberrant ToM is associated with cognitive-perceptual, but not interpersonal or disorganized manifestations. To date, only one study has assessed the associations between ToM performance and schizotypal trait dimensions in community adolescents. This has replicated previous findings in adult samples suggesting that poor ToM performance is associated with increased scores in the cognitive-perceptual dimension of schizotypy.

Interestingly, in line with studies in people suffering with clinical psychosis, emerging evidence suggest significant associations between a hypermentalizing ToM pattern (i.e. overattributing intentions in others) and a history of psychotic-like experiences in samples of older children from the general population.

As it pertains to psychological processes involving the understanding of the self, research has shown that patients with schizophrenia, individuals at clinical high-risk for the illness (CHR), as well as non-clinical individuals who self-report schizotypal personality traits and psychotic-like experiences display an externalizing tendency towards: (a) misattributing self-generated imagined mental events as overtly enacted (Docherty, 2012; Franck et al., 2000; Gawęda et al., 2012; Henquet et al., 2005; Humpston et al., 2017; Peters et al., 2007); and (b) misattributing the source of self-generated actions to external agents (Bentall et al., 1991; Brébion et al., 2000; Johns et al., 2010; Johns et al., 2006; Vermissen et al 2007; Allen et al., 2006; Larøi et al., 2004). Furthermore, these patterns of externalizing misattributions are found to be particularly pronounced among patients reporting hallucinations (Bentall et al., 1991; Brébion et al., 2000) and non-clinical adults experiencing hallucination-like phenomena (Larøi et al., 2004) Importantly however, although different types of self- and reality-monitoring misattributions have also been identified in studies of adolescents who are at neurogenetic risk for psychosis (Debbané et al., 2008; 2010), no study to date has examined the associations between these self-referential processes and schizotypal dimensions in typically developing adolescents (see Tables 1.4. and 1.5. for a description of the self- and reality-monitoring studies reviewed in this chapter).

Study	Sample	Comparison group	Task used	Task type
Clinical Psychosis				
Keefe et al., (1999)	Patients with schizophrenia (n = 28)	healthy controls (n = 19)	Say/imagine	Verbal
Franck et al., (2000)	Patients with hallucinations (n = 17)	healthy controls (n = 17)	Say/imagine	Verbal
Docherty, (2012)	schizophrenia/schizoaffective disorder patients (n = 110)	healthy controls (n = 23)	Say/imagine	Verbal
Henquet et al., (2005)	Patients with schizophrenia (n = 15)	psychiatric controls (n = 15)	Say/imagine	Verbal
CHR				
Gawęda et al., (2017)	Individuals meeting UHR criteria (n = 36) and FEP patients (n = 25)	Healthy controls (n = 33)	Preform/imagine	Action
Schizotypy				
Peters et al., (2007)	undergraduate students (n = 67)	N/A	Perform/imagine	Action
Humpston et al., (2017)	Non-clinical adults (n = 102)	N/A	Perform/imagine	Action
Debbané et al., (2010)	Adolescents with 22q11.2 deletion syndrome (n = 20)	IQ- and age-matched controls (n = 19); age-matched controls (n = 19)	Say/imagine	Verbal

Table 1.4. Characteristics of the self-monitoring studies reviewed in this chapter

Abbreviations: CHR, Clinical High Risk for psychosis; UHR, Ultra High Risk for psychosis; FEP, First episode psychosis

Study	Sample	Comparison group	Task used	Task type
Clinical Psychosis				
Bental et al., (1991)	Patients with hallucinations (n = 22); patients with delusions and no hallucinations (n = 16)	healthy controls (n = 16)	source memory	Verbal
Brébion et al., (2000)	Inpatients with schizophrenia (n = 40)	healthy controls (n = 40)	source memory	Verbal
Morrison and Haddock, (1997)	Patients with auditory hallucinations (n = 15)	Psychiatric controls (n = 15); healthy controls (n = 15)	source memory	Verbal
Johns et al., (2001)	Patients with hallucinations but no delusions (n = 10); Patients with delusions but no hallucinations (n = 8)	Non-clinical controls (n = 20)	Action-simulation	Verbal
Franck et al., (2001)	Patients with schizophrenia and delusions of reference (n = 6); patients with schizophrenia without delusions of reference	Non-clinical controls (n = 29)	Action-simulation	Motor
Versmissen et al., (2007)	Patients with schizophrenia (n = 37)	Non-clinical adults with high schizotypy (n = 44); first-degree relatives of patients (n = 41); healthy controls (n = 49)	Action-simulation	Motor
Homes et al., (2011)	Patients with schizophrenia (n = 42)	First-degree siblings of patients with schizophrenia (n = 32); non-clinical controls (n = 49)	Action-simulation	Motor
Knoblich et al., (2004)	Patients with schizophrenia (n = 27)	Non-clinical controls (n = 23)	Action-simulation	Motor
Blakemore et al., (2000)	Patients with auditory hallucinations and/or passivity experiences (n = 17)	Non-hallucination/non-passivity experiences psychiatric controls (n = 24); non-clinical controls (n = 15)	Tactile stimulus	Motor
CHR				
Johns et al., (2010)	Adults meeting UHR criteria (n = 31)	Age-matched controls (n = 31)	Action-simulation	Verbal
Schizotypy				
Versmissen et al., (2007)	Non-clinical adults with high schizotypy (n = 44)	Patients with schizophrenia (n = 37); first-degree relatives of patients (n = 41); healthy controls (n = 49)	Action-simulation	Motor
Homes et al., (2011)	First-degree siblings of patients with schizophrenia (n = 32)	Patients with schizophrenia (n = 42); non-clinical controls (n = 49)	Action-simulation	Motor

Larøi et al., (2004)	Non-clinical adults scoring high on hallucination proneness (n = 25)	Non-clinical adults scoring low on hallucination proneness (n =25)	Source memory	Verbal
Allen et al., (2006)	Non-clinical adults (n = 57)	N/A	Action-simulation	Verbal
Larøi et al., (2005)	Non-clinical adults scoring high on hallucination proneness (n = 16)	Non-clinical adults scoring low on hallucination proneness (n =16)	Source memory	Motor
Alderson-Day et al., (2019)	Non-clinical young adults (n = 76)	N/A	Source memory	Verbal
Debbané et al., (2008)	Adolescents with 22q11.2 DS (n = 18)	Non-clinical age- and IQ-matched controls (n = 17); Age-matched controls (n = 17)	Source memory	Motor
Debbané et al., (2009)	Non-clinical adolescents (n = 163)	N/A	Source memory	Motor

Table 1.5. Characteristics of reality-monitoring studies reviewed in this chapter

Abbreviations: CHR, Clinical High Risk for psychosis; UHR, Ultra High Risk for psychosis

Overall, the review of the literature suggests that disturbances in higher order processes pertaining to the understanding of the self and others are present across the continuum of psychosis expression and may contribute to the developmental unfolding of the illness in those who are at increased clinical risk. Most importantly perhaps, increasing evidence indicate that subtle disruptions in these mentalizing processes relate to premorbid risk for psychosis in the context of schizotypal trait expression, prior to the development of clinically-relevant symptoms or clinical risk states. From a clinical standpoint this suggests that psychotherapeutic approaches targeting disruptions in the understanding of oneself and others may be applied preventatively in order to inhibit the trajectory of emerging psychosis in people who present with non-clinical schizotypal manifestations.

Critically however, the review of the literature also highlights the relative lack of studies in adolescent populations. This is important as adolescence represents a key period for the emergence of schizotypal trait manifestations, but also for the developmental elaboration of higher order cognitive processes pertaining to mentalizing. Thus, from a dynamic, developmental point of view it can be hypothesised that cognitive, perceptual and interpersonal aberrations arising in the context of an underlying schizotypal trait liability may contribute to disrupt the normative development of mentalizing processes during adolescence and impair the capacity for self and other understanding (Debbané & Barrantes-Vidal, 2014). Difficulties in recognising inner states in oneself and others may in turn lead to the exacerbation of the underlying trait liability in a transactional process that potentiates psychosis vulnerability from adolescence to young adulthood and increases the risk of transitioning to a first episode of psychosis (Debbané, Salaminios et al., 2016; Debbané & Barrantes-Vidal, 2014). Investigating the nature of associations between emerging schizotypal traits and mentalizing processes in typically developing adolescents can therefore increase our understanding regarding the earliest stages of a pathogenic developmental processes that distally relates to the expression of psychotic states.

Importantly, even among the relatively small number of studies that have directly assessed the relationship between schizotypal traits and mentalizing processes in adult populations, a number of empirical and conceptual gaps can also be identified. First the preponderance of research in adult samples has focused on cognitive-perceptual manifestations, either by using measures that assess hallucination and delusion-like phenomena, or by undertaking group comparisons according to the presence or absence of these symptoms. At present, it remains unclear whether schizotypal traits that specifically manifest within the interpersonal domain may also relate to mentalizing difficulties. This is surprising given that (1) the development of mentalizing processes commonly unfolds against

the background of close interpersonal relationships (Fonagy et al., 2002); and (2) schizotypal traits that directly manifest within the interpersonal domain have been specifically associated with an elevated risk for the development of schizophrenia-spectrum disorders (Debbané et al., 2014; Flückiger et al., 2016).

Second, the assessment of mentalizing in psychosis-research has traditionally relied on the use of experimental tasks that measure participants' performance on specific domains of mental state understanding (e.g. ToM; reality-monitoring). While, the use of experimental paradigms carries the considerable advantage of providing an objective measurement of mentalizing dimensions, these are limited in assessing only singular aspects of the broader mentalizing construct, thus their application is not feasible in the context of real-world clinical settings (i.e. as part of routine outcome monitoring in clinical services) (Boldrini et al., 2020). In addition to task-based measures, a small amount of research has examined mentalizing in people diagnosed with psychosis and individuals exhibiting prodromal symptoms by using interview-based methodologies, yielding interesting results. For instance, Braehler and Schwannauer (2012) found that the level of mentalizing capacity, as assessed through the Reflective Functioning Scale (RFS, Fonagy et al., 1998), moderated functional outcomes in adolescents recovering from a first episode of psychosis. In a more recent study using the RFS, Boldrini et al (2020) reported that mentalizing difficulties prospectively predicted transition to clinical psychosis in a sample of prodromal adolescents and young adults. Although narrative-based methodologies, such as the ones used in these studies, offer the opportunity to undertake complex multidimensional assessments of mentalizing, thus could be applied in the context of clinical settings, they are often lengthy and require expert training in their administration and rating. An easier-to-administer alternative method to assess mentalizing is through self-report. However, self-report measures of mentalizing have seldomly been used in psychosis research. This is primarily due to the potential confounding effects that illness-specific factors (e.g.

cognitive decline, lack of insight, side effects of chronic pharmacological treatment) may bring on one's subjective report of their mentalizing capacities (Fonagy et al., 2016). Importantly though, the assessment of mentalizing through self-report may be feasible in non-clinical samples of individuals who exhibit schizotypal manifestations, thus are not affected by the long-term effects of chronic psychosis. To date, no study has examined the associations between schizotypal trait manifestations and self-reported mentalizing difficulties.

Finally, most experimental studies that have examined the relationship between mentalizing abilities and psychotic phenomena have done so either through the use of ToM tasks, which only capture the capacity to make inferences about other peoples' mental states (e.g. facial affect recognition, prosody affect recognition, false belief understanding, sarcasm perception, etc.), or by assessing disruptions in the capacity to recognise and monitor one's own self-generated mental events (i.e. mentalizing about the self) (Allen et al., 2007; Blakemore et al., 2003). Yet little is known at present about the developmental links between these two processes. This is important as contemporary models of psychosis development suggest that impairments in self-referential processing during the premorbid period of psychosis expression may contribute to the abnormal explanations of social experience commonly observed during the prodromal and clinical stages of the illness (Brent and Fonagy, 2014; Debbané, Salaminios et al., 2016). Within this context, assessing mentalizing functions that involve inferences about the mental states of others, such as ToM, while still in the process of development during adolescence and young adulthood, offers the opportunity to examine whether these are critically underpinned by more basic metacognitive abilities pertaining to the monitoring of the self.

1.8. Overview of the empirical studies of the thesis

To address the gaps in the previous literature, the empirical studies undertaken will assess mentalizing processes in samples of typically developing adolescents. These will utilise novel experimental tasks and self-report measures specifically geared to assess psychological processes pertaining to the understanding of one's own and other peoples' mental events. The main aim of the thesis will be to assess the nature of associations linking the expression of schizotypal traits to mentalizing difficulties in adolescence. In contrast to the majority of previous studies that have focused solely on positive schizotypy, the empirical studies of the current thesis will undertake a multidimensional assessment of schizotypal traits that will encompass manifestations across the perceptual, interpersonal and cognitive domains. To better understand the process of adolescent mentalizing, particularly in relation to the monitoring of self-generated mental events, the current thesis will also seek to explore the effects that developmental (i.e. age), cognitive (i.e. cognitive effort) and affect-based factors (emotional valence) may exert on psychological processes pertaining to the understanding of the self (i.e. self- and reality-monitoring). Finally, to bridge the gap between the higher order psychological processes involves in self- and other-understanding in the study of schizotypy, the current thesis will also explore the prospective effects of self-monitoring difficulties, as well as their interactions with schizotypal trait dimensions in adolescence, on the capacity to attribute mental states in others (i.e. ToM).

Study 1: **Associations Between Schizotypal Personality Features, Mentalizing Difficulties and Thought Problems in a Sample of Community Adolescents** (Chapter 2).

Increasing evidence suggest that clinical psychosis is distally linked with schizotypal trait expression and more proximally with the breakdown of higher-order cognitive processes pertaining to mentalizing. Importantly, subtle mentalizing difficulties have been identified

among non-clinical individuals in the context of high psychometric schizotypy, prior to the development of clinical symptoms, suggesting an early pathway towards illness expression. However, little is known at present about the relationship between specific schizotypal trait features and mentalizing during the critical developmental period of adolescence. Furthermore, no study to date has investigated the mediational role of mentalizing on the relationship between schizotypy and clinically-relevant symptoms. Study 1 will investigate, for the first time, the associations between schizotypy and self-reported mentalizing in adolescence, as well as the extent to which mentalizing difficulties mediate the relationship between schizotypy and thought problems relevant for psychosis. To this end study 1 will adopt a clinically-relevant approach by: (a) investigating schizotypy at the individual trait-feature level (rather than measuring total dimension scores); (b) using an easy-to-administer self-report measure of general mentalizing abilities that can be utilised within clinical settings for screening and routine outcome evaluation purposes (Reflective Functioning Scale, RFQ); and (3) assessing thought problems previously linked to proximal risk for psychotic disorders.

Study 2: Self-Monitoring for Speech and its Links to Age, Cognitive Effort, Schizotypal Trait Expression and Impulsivity During Adolescence (Chapter 3).

Disruptions in self-monitoring - the capacity to discriminate between self-generated real and imagined events - represent key factors associated with schizophrenia spectrum disorders. In addition, increasing evidence suggest that self-monitoring difficulties are already present during the early stages of emerging psychosis, among help-seeking individuals manifesting CHR state manifestations and in community samples of adults who report high psychometric schizotypy. No study to date has investigated self-monitoring performance or its associations with schizotypal personality traits among typically developing adolescents. While the first study of the current thesis focused on the assessment of general mentalizing skills in

adolescence through self-report, the second study will narrow the focus of investigation to the empirical assessment of psychological abilities involved in the understanding of the self. Specifically, study 2 will test self-monitoring by using an experimental task that assesses the capacity to discriminate between self-generated overt and silent speech. The aims of the study are to test the effects of age and cognitive effort on self-monitoring performance in adolescence, as well as its associations with trait dimensions pertaining to schizotypy and impulsivity.

Study 3: Reality-monitoring in adolescence and its links to age, emotional valence and schizotypal trait expression (Chapter 4).

A large body of research has shown that patients with schizophrenia and individuals at clinical high-risk for the illness (CHR) display a tendency towards misattributing the source of self-generated items to external agents. This pattern of externalizing reality-monitoring misattributions has also been identified among adults reporting non-clinical hallucination and delusion-like phenomena. At present, no study has directly assessed reality-monitoring or its associations with schizotypal trait dimensions during the critical developmental window spanning from adolescence to young adulthood. Following the second study, which focused on self-oriented processes involving the capacity to discriminate between different sources of self-generated information (i.e. what one imagined saying from what one overtly said), the third study of the thesis will extend the focus of investigation to the relationship between schizotypal dimensions and the capacity to distinguish between self-generated and other-produced information. Specifically, study 3 will examine reality-monitoring performance in community adolescents and young adults by using a task that assesses the capacity to discriminate in memory between self-generated and experimenter-produced speech. The aims of the study are to assess the effects of age and emotional valence on reality-monitoring for speech in

adolescence and young adulthood, as well as its associations with personality dimensions pertaining to schizotypy.

Study 4: Prospective effects of schizotypal trait expression and self-monitoring on different types of ToM dysfunction during adolescence and young adulthood (Chapter 5).

Disruptions in ToM have been consistently associated with psychotic phenomena, both the state and trait levels of their expression. To date only a small number of cross-sectional studies have assessed ToM performance in adolescents and young adults in the context trait risk for psychosis, suggesting associations with cognitive-perceptual manifestations. Importantly however, no study to date has examined prospective associations between schizotypal trait dimensions and ToM between adolescence and young adulthood. In a similar vein, no study has assessed whether different schizotypal dimensions are associated with different types of ToM dysfunction (i.e. hypomentalizing vs hypermentalizing). Following the previous studies of the current thesis that focused on thinking processes sustaining the awareness of the self, the final study will examine prospective associations between schizotypal trait dimensions and the awareness of inner states in others. More specifically, study 4 will test ToM performance in community adolescents and young adults by using the a novel video-based task that offers the unique advantage of independently assessing the use of aberrant ToM strategies, such as hypermentalizing and hypomentalizing, for both cognitions and affects. Furthermore, the study will attempt to bridge the divide between the higher order psychological processes underpinning self- and other-understanding by testing whether disruptions in metacognitive abilities sustaining the awareness of the self (i.e. self-monitoring) will prospectively predict dysfunctions in higher-order cognitive processes pertaining to the awareness of others (i.e. ToM). The main aims of the study are to prospectively assess the effects of schizotypal trait

dimensions, self-monitoring difficulties and their interactions during adolescence on different types of ToM dysfunction at 5-year follow-up.

1.8.1. Contributions to research

The current thesis has drawn upon previously collected data from samples of typically-developing adolescents and young adults that participated in large-scale research on developmental psychopathology, led by Dr Debbané at the University of Geneva. Although I was not directly involved in the collection of participant data, I had regular meetings and liaised closely with the researchers involved in data collection. Furthermore, I was involved in designing and planning the collection of follow-up data, which were specifically collected for study 4 of the current thesis (see chapter 5). In addition, I was responsible, under the supervision of Dr Debbané, to design the studies presented in the current thesis, which included identifying relevant hypotheses, developing the conceptual rationale and choosing the statistical methods to test these. Regarding the results of the current thesis, I independently undertook all the statistical analyses of the data used. Finally, I was the first author of two peer-reviewed journal publications that present the findings from studies 1 and 2 of the thesis (see chapters 2 and 3) and joint first author (with Dr Debbané) of one journal publication (Debbané, Salaminios et al., 2016) and one scientific book chapter (Salaminios and Debbané, 2021) on the topic of mentalizing in emerging psychosis.

Chapter 2

Associations Between Schizotypal Personality Features, Mentalizing Difficulties and Thought Problems in a Sample of Community Adolescents[1]

2.1 Introduction

Contemporary research suggests that clinical psychosis is distally linked with schizotypal trait expression (Chapman et al., 1994; Kwapil et al., 2013), and more proximally with the breakdown of higher-order cognitive processes pertaining to mentalizing (Kim et al., 2011) – the capacity to understand the intentional mental states driving one's own and others' behaviours (Fonagy & Target, 1996). As discussed in Chapter 1 mentalizing represents a multifaceted construct that encompasses a number of processes involved in mental state understanding, such as theory of mind (ToM: inferring beliefs from others' behaviours), empathy (understanding the emotional states of others), mindfulness (emotional self-awareness), and self-monitoring (cognitive self-awareness) (Choi-Kain & Gunderson, 2008; Debbané et al., 2016; Fonagy & Luyten, 2009). Good mentalizing entails the integration of these processes to construe a sensitive understanding of the internal states (i.e. thoughts, feelings) underpinning one's own and others' intentions (Fonagy et al., 2002; Fonagy & Luyten, 2009). As such, mentalizing enables individuals to form representational models of human behaviour in order to sustain a coherent sense of self and attenuate experiences of confusion within interpersonal relationships (Fonagy et al., 2002).

Difficulties in utilizing mental state information to understand oneself and others are increasingly regarded as key factors for the development and maintenance of psychotic

---

[1] The current chapter has been published in Salaminios, G., Morosan, L., Toffel, E., Tanzer, M., Eliez, S., Badoud, D., Armando, M., Debbané, M. (2020) Associations Between Schizotypal Personality Features, Mentalizing Difficulties and Thought Problems in a Sample of Community Adolescents. *Early Intervention in Psychiatry*

manifestations (Brent & Fonagy, 2014; van Os et al., 2010). Meta-analytic investigations indicate that patients with established schizophrenia, first episode sufferers, as well as help-seeking individuals at clinical high-risk for psychosis (CHR), perform poorly in multiple domains of mentalizing, such as ToM, self-monitoring, and emotion recognition (Bora & Pantelis, 2013; Brüne, 2005; Sprong et al., 2007). Overall, current evidence suggest the presence of a generalized impairment in mentalizing processes across schizophrenia-spectrum disorders.

Another line of research indicates that subtle mentalizing difficulties are present among non-clinical samples in the context of schizotypal personality features, prior to the development of clinical manifestations, suggesting a pathway towards illness expression (Langdon & Coltheart, 1999; Pickup, 2006). Schizotypal traits signal an increased liability to develop schizophrenia (Barrantes-Vidal et al., 2015; Chapman et al., 1994; Kwapil et al., 2013) and most psychometric analyses examining their factorial structure have consistently identified three main dimensions: the cognitive-perceptual (positive schizotypy: hallucination, delusion-like phenomena), interpersonal (social anxiety, constricted affect), and disorganization dimensions (odd behaviours, odd speech) (Fonseca-Pedrero et al., 2018; Raine, 2006). At present, only a limited number of correlational studies have directly analysed the associations between schizotypal trait dimensions and mentalizing difficulties in community samples. These have reported that high scores on the cognitive-perceptual dimension are associated with reduced ToM scores (Langdon & Coltheart, 1999; Pickup, 2006). Although these findings are suggestive of a relationship between trait-vulnerability for psychotic disorders and mentalizing difficulties, they involve adult participants either within or beyond the critical period of illness onset. From an early intervention standpoint, important insights can be gained by examining mentalizing in the context of schizotypal trait dimensions during earlier developmental stages, in childhood and adolescence.

Schizotypal manifestations emerge early in life (Debbané & Barrantes-Vidal, 2014) and adolescence represents a key period during which they can become clinically relevant (Gooding et al., 2005). Developmentally, the maturation and elaboration of mentalizing processes has been shown to continue throughout adolescence and into early adulthood (Choudhury et al., 2006; Dumontheil et al., 2010; Vetter et al., 2013). As such, cognitive and interpersonal aberrations arising in the context of schizotypal personality traits, may disrupt the normative development of mentalizing processes during adolescence and impair the capacity for mental state understanding (Debbané & Barrantes-Vidal, 2014). Indeed, recent evidence indicate that young people who display attenuated psychotic manifestations demonstrate reduced trajectories for common age-related improvements in ToM compared to healthy controls (Davidson et al., 2018). Because the development of mentalizing commonly unfolds against the background of close interpersonal relationships (Fonagy et al., 2002), one possibility is that the expression of schizotypal traits that impede interpersonal communication during adolescence may negatively impact on the capacity to understand mental states. However, little is known at present about the nature of associations linking specific schizotypal personality features to mentalizing difficulties during adolescence.

Furthermore, emerging research suggests that the effects of schizotypal traits on the development of clinically-relevant symptoms may depend upon the level of mentalizing dysfunction. For instance, Bartels-Velthuis et al. (2011), reported that in sample of young adolescents who experienced non-clinical hallucinations within the previous 5 years, the risk for the development of secondary delusional ideation was higher for those with lower ToM scores. These findings suggest that mentalizing dysfunction represents an independent factor that may interact with schizotypal traits during adolescence to moderate their effects on the development of psychotic symptoms. Importantly however, evidence suggest that schizotypal traits and mentalizing difficulties are not independent from each other. As discussed above, a

number of studies have reported significant associations between schizotypal traits and mentalizing difficulties among both non-clinical adolescent and adult samples. Based on these data, an alternative hypothesis that remains empirically unexplored pertains to the potential mediating role of mentalizing on the relation between schizotypal personality traits and the experience of clinically-relevant symptoms in adolescence. According to this hypothesis, the expression of schizotypal personality traits during adolescence may contribute to impair the developmental capacity to understand mental states, and these impairments may in turn contribute to the experience of psychotic symptoms (Debbané & Barrantes-Vidal, 2014; Salaminios et al., 2021). In other words, mentalizing dysfunction may represent a pathway through which schizotypal personality traits, at least in part, augment psychosis risk. While previous research has shown the mentalizing difficulties mediate the effects of interpersonal factors (i.e. attachment insecurity) on the experience of thought problems that lie in a continuum with clinically-relevant symptoms during adolescence (Hart et al., 2017), no study to date has examined whether mentalizing also mediates the relation between schizotypal personality features and thought problems in adolescence. Research on adolescent schizotypy can elucidate our understanding regarding the early course and significance of mentalizing dysfunction in the premorbid stages of schizophrenia-spectrum illnesses, with important implications for early prevention treatment.

Mentalizing in psychosis-research has been traditionally measured using experimental-tasks that assess specific domains of mental state understanding (e.g. ToM). While these are effective in empirically assessing singular components of mentalizing they are often time- and labour-consuming, thus not suitable for use in context of clinical settings or large-scale research studies. Recently, Fonagy et al. (2016) developed the Reflective Functioning Questionnaire (RFQ), an easy-to-administer, self-report screening measure of general mentalizing abilities. The RFQ assesses participant's self-reported *certainty* and *uncertainty* about mental states,

reflecting how confident versus how doubtful one is in utilizing mental state information, such as thoughts and feelings, to form representations of their own and others' behaviours. Because of its brief nature, the RFQ is a suitable assessment tool for the purposes of routine outcome evaluation in clinical settings and clinical trials and its use has been validated for adolescents from the general population (Badoud et al., 2015). In addition, previous studies suggest that the RFQ correlates with measures of mindfulness, perspective-taking and empathy (Fonagy et al., 2016), thus appears to be effective in capturing different facets of the wider mentalizing construct.

The present study seeks to, for the first time, utilize the RFQ in the study of adolescent schizotypy in order to: (1) identify specific features across all three schizotypal dimensions that account for RFQ-measured mentalizing difficulties in adolescence; (2) assess the effects of schizotypal personality features and RFQ-measured mentalizing on thought problems relevant for psychosis; and (3) examine whether RFQ-measured mentalizing mediates the relation between schizotypal personality features and thought problems.

Given that the normative elaboration of mentalizing relies upon interpersonal interactions within close relationships (Fonagy et al., 2002; Gergely & Watson, 1999; Uddin et al., 2007), it is hypothesised that schizotypal features that impede interpersonal communication during adolescence will be linked to mentalizing difficulties. Specifically, it is expected that social anxiety, constricted affect and lack of close friends (interpersonal), as well as suspiciousness (cognitive-perceptual) and odd speech (disorganization) will be associated with high uncertainty and low certainty in understanding mental states. Furthermore, on the basis of previous research (Debbané et al., 2014; Hart et al., 2017) it is hypothesised that both schizotypal personality features and mentalizing difficulties will be associated with self-rated thought problems. Finally, given that mentalizing dysfunction has been proposed as a potential pathway through which schizotypal features relate to psychosis vulnerability (Bartels-Velthuis

et al., 2011; Debbané et al., 2016), it is expected that mentalizing difficulties (high uncertainty and low certainty in mental states) will, at least in part, mediate the effects of schizotypal personality features on thought problems in our community adolescent sample.

2.2 Methods

2.2.1 Participants and procedure

A convenience sample of 120 community adolescents was recruited through written advertisements in public schools in the city of Geneva, Switzerland. No prior power analysis was performed to determine the size of the sample. Inclusion criteria were age (12-19 years) and fluency in French. Participants were screened for cognitive impairment, and those with a standard score below 7 in the Block Design subtest of the Wechsler Intelligence Scale (Wechsler, 1955) were excluded from the analyses (n = 15). The final sample consisted of 105 community adolescents (53 females, 52 males, M_{age} = 15.72). None of the participants suffered from past/present psychiatric, neurological or neurogenetic disorders. Written informed consent was obtained from all participants and legal guardians of those under 18 years of age. Ethical approval for the study was granted by the Swiss Ethics Committee on research involving humans (number: 2018-00251).

2.2.2 Measures

The following battery of self-report questionnaires was individually administered in random order to participants under the supervision of a trained research assistant to make sure participants understood the items.

The Schizotypal Personality Questionnaire (SPQ, French version, Dumas et al., 2000) is designed to measure schizotypal traits subjectively experienced as common aspects of one's personality functioning. The measure yields three dimension scores and nine subscale scores: cognitive–perceptual (unusual perceptual experiences, ideas of reference, suspiciousness, odd

beliefs or magical thinking), interpersonal (social anxiety, constricted affect, lack of close friends), and disorganization (odd speech, odd behaviour) (see appendix 1). The French version of the SPQ has shown good reliability (Cronbach's alpha = 0.91) (Dumas et al., 2000) and has been validated for use with francophone adolescents (Badoud et al., 2011). The SPQ in the present study also showed good reliability (Cronbach's alpha = 0.87).

The *Reflective Functioning Questionnaire*, brief version (RFQ, French version, Badoud et al., 2015) includes eight items (see appendix 2), which are self-rated on a seven-point scale from 1 (strongly disagree) to 7 (strongly agree). RFQ items aim capture self-reported mentalizing abilities by measuring the degree of certainty and uncertainty with which individuals utilize mental state information to understand their own (e.g. items 2, 6, 7) and others' behaviour (e.g. item 1), including during moments of increased affective arousal (e.g. items 3-5, 8). The measure is comprised of two subscales. The *uncertainty about mental states* subscale (RFQu; items 2, 4, 5, 6, 7, 8 in appendix 2) focuses on the extent to which individuals agree with statements such as "Other people's thoughts are a mystery to me" and "Strong feelings often cloud my thinking". RFQu item scores are recoded (1 = 0; 2 = 0; 3 = 0; 4 = 0; 5 = 1; 6 = 2; 7 = 3), so that higher scores reflect poor usage of mental state information and a stance characterized by a lack of knowledge about mental states, while lower scores reflect acknowledgement of the opaque nature of mental states. The *certainty about mental states* subscale (RFQc; items 2, 4, 5, 6, 1, 3 in appendix 2.) focuses on the extent to which individuals disagree with statements such as "I don't always know why I do what I do". RFQc item scores are recoded (1 = 3; 2 = 2; 3 = 1; 4 = 0; 5 = 0; 6= 0; 7 = 0) so that higher scores on this scale reflect better usage of mental state information and adaptive levels of certainty about mental states. The measure has previously been validated for French-speaking adolescents (Badoud et al., 2015), showing satisfactory internal consistency for both the RFQu (Cronbach's alpha = 0.68) and RFQc scales (Cronbach's alpha = 0.74). In the present study these values were

comparable to the sample of Badoud et al., (2015) with Cronbach's alpha values of 0.68 for the RFQu scale and 0.73 for the RFQc scale.

The *Youth Self-Report* (YSR; Achenbach, 1991) measures self-reported psychopathology in the past 6 months among adolescents aged 11-18 years. The *Adult Self-report* (ASR; Achenbach, 1991) assesses psychopathology in adults. It has comparable items to the YSR and the same range of scores across its scales (see appendix 3). In line with recent studies (Hart et al., 2017), the *thought problems* subscale of the YSR/ASR was used to assess cross-cutting, rather than disorder-specific, symptoms that lie in a continuum with clinical manifestations relevant for psychosis, including cognitive-perceptual aberrations, as well as strange thoughts and behaviours. Assessing psychosis-relevant symptoms continuously, rather than focusing solely on specific symptom-subgroups, addresses the need to identify cross-cutting factors that contribute to early risk for psychosis in adolescence. Indeed, longitudinal findings suggest that the *thought problems* subscale captures symptomatic behaviours relevant to proximal/prodromal risk for psychosis (Simeonova et al., 2014). Prior psychometric evaluations have shown that the YSR/ASR scales demonstrate adequate validity and reliability, with Cronbach's alpha value of 0.78 for the thought problems subscale (Achenbach, 1991). Because symptoms of anxiety and depression have been shown to influence mentalizing inferences (Bateman & Fonagy, 2004) the *anxious/depressed* subscale of the YSR/ASR was used as a control variable in the analyses.

2.2.3 Statistical analyses

The main statistical analyses were performed in SPSS version 24. Spearman's correlation coefficients were calculated to examine associations between SPQ and RFQ subscale scores (Bonferroni corrected alpha = 0.001). Two multiple linear regression models were ran to assess

the effects of SPQ subscales on RFQ scale scores (RFQu and RFQc). In the first model, all SPQ subscales were entered together as independent variables and RFQu was entered as the dependent variable. In the second model, all SPQ subscales were entered together as independent variables and RFQc was entered as the dependent variable. Age and the YSR/ASR anxious/depressed subscale were also entered in each model to account for their effects. Multicollinearity was tested using variance inflation factors (VIF) and tolerance statistics (VIF > 5 and/or tolerance < 0.2 suggest multicollinearity).

Two multiple linear regression models, controlling for age, were also ran to explore the effects of (a) SPQ subscale scores and (b) RFQ scale scores, on the level of thought problems (YSR/ASR thought problems subscale T-score).

Given that no prior power calculations were performed to determine the sample size needed to detect significant findings and in order to explore whether the results of the regressions conducted in the main analyses represent true effects as opposed to non-conclusive findings, JASP version 0.9.2 was used to obtain simple Bayesian statistics of the regression analyses ran (Van den Bergh et al., 2021). In contrast to standard hypothesis testing based on P values, Bayesian regression analyses offer a direct measure of the strength of the evidence both in favour and against the study hypotheses, thus supporting the interpretation of results and reducing the probability of type II errors. More specifically, rather than basing inferences on a single model, Bayesian regression analysis retains all possible models for inference and quantifies the importance of individual predictors by providing a Bayes factor ($BF_{inclusion}$), which signals the factor by which the prior probabilities of including a predictor in a model (calculated by dividing the number of models that contain the predictor from the number of all possible models) increase or decrease after observing the data (i.e. after averaging the predictive odds in favour of the models that contain the individual predictor). $BF_{inclusion}$ values below 1.00 indicate the factor by which the data have decreased the prior odds for including a

predictor in a model, while values above 1.00 show the factor by which the prior odds for including a predictor in a model have increased based on the data.

Finally, Hayes's PROCESS macro (Hayes and Preacher, 2013) for SPSS version 24 was used to test the prediction that mentalizing mediates the relationship between schizotypal traits and subthreshold psychotic symptoms. The YSR/ASR thought problems scale was entered as the outcome variable, SPQ subscales were entered as the independent variables and RFQ scales as the mediators in each model, with age as a control variable in the analyses. The PROCESS macro uses non-parametric bootstrapping, which involves random resampling of observations with replacement to obtain confidence intervals for indirect effects (bootstrap confidence intervals were based on 10.000 samples). Effects are considered significant if the confidence interval does not contain zero. Mediation is present if the indirect effect (a and b) of the independent variable (SPQ subscale) on the outcome variable (YSR/ASR thought problems) through the mediator (RFQ scale) is significant and the direct effect of the dependent variable on the outcome variable, while accounting for the mediator (c'), is smaller than the total effect (c) (figure 2.1.).
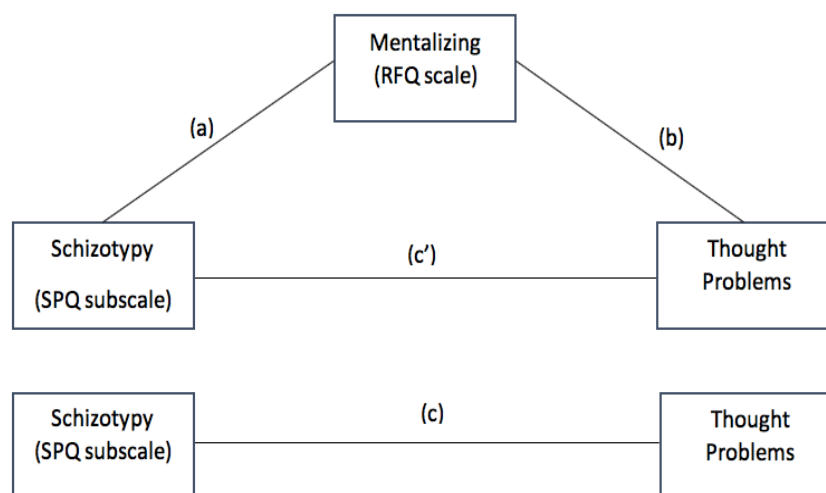
Figure 2.1: Mediational model with schizotypy as independent variable, mentalizing as mediator and thought problems as dependent variable

2.3. Results

2.3.1 Descriptive results

Table 2.1. presents the descriptive results for the major variables included in the analyses.

	Mean	Range	SD	Skewness	Kurtosis
Mentalizing					
RFQu	3.02	0-11	2.55	0.74	-0.06
RFQc	6.36	0-18	4.04	0.67	0.07
SPQ Cognitive-perceptual dimension					
SPQ Unusual perceptual experiences	2.24	0-8	2.20	0.83	-0.42
SPQ Ideas of reference	2.32	0-8	2.23	0.89	-0.15
SPQ Suspiciousness	1.99	0-7	1.83	0.90	-0.08
SPQ Odd beliefs	1.89	0-7	1.78	0.86	-0.08
SPQ Interpersonal dimension					
SPQ Social anxiety	2.86	0-8	2.27	0.48	-0.78
SPQ Constricted affect	2.00	0-6	1.50	0.43	-0.65
SPQ Lack of close friends	1.72	0-9	1.67	1.48	1.98
SPQ Disorganised dimension					
SPQ Odd speech	3.96	0-9	2.65	0.21	-1.05
SPQ Odd behaviour	1.65	0-6	1.83	0.98	-0.29
Psychological Symptoms					
YSR/ASR Thought problems	58.68	50-95	8.22	1.16	1.42
YSR/ASR Anxious/depressed	56.31	50-94	7.78	1.62	2.22

Table 2.1. Means, ranges and standard deviations for each major variable in the sample

Abbreviations: RFQu, RFQ Uncertainty about Mental States; RFQc, RFQ Certainty about Mental States; SPQ, Schizotypal Personality Questionnaire; YSR/ASR, Youth and Adult Self Report

Kurtosis and skewness of the distributions for the RFQ scales and SPQ subscales were all within the acceptable range (i.e. all values between -2.00 and 2.00) and no outliers were identified. Furthermore, analyses of the SPQ distributions showed that the proportion of participants reporting higher schizotypal trait scores (1.5 standard deviations above the normed means) was comparable across the majority of SPQ subscales (12.38% for unusual perceptual

experiences [n = 13]; 12.38% for ideas of reference [n = 13]; 13.33% for suspiciousness [n = 14]; 12.38% for odd beliefs [n = 13]; 13.33% for social anxiety [n = 14]; 11.43% for odd speech [n = 12] and 11.43% for odd behaviour [n = 11]). It must be noted that the proportions of participants with scores 1.5 standard deviations above the normed means on the constricted affect (7.62%, n = 8) and lack of close friends (7.62%, n = 8) was lower compared to the other SPQ subscales. In terms of mentalizing, 11.43% of participants (n = 12) scored 1.5 standard deviations above the normed means on the RFQu scale compared to 7.62% (n = 8) on the RFQc scale.

2.3.2 Assessing the relationships between schizotypal trait dimensions and mentalizing difficulties in adolescence

Table 2.2. presents the Spearman's correlations between SPQ and RFQ subscale scores.

	1	2	3	4	5	6	7	8	9	10	11
1. RFQu	-	-0.42**	0.21	0.36**	0.34**	0.38**	0.34**	0.20	0.40**	0.18	0.30
2. RFQc		-	-0.16	-0.18	-0.23	-0.24	-0.30	-0.24	-0.34**	-0.05	-0.14
3. SPQ Unusual perceptual experiences			-	0.53**	0.48**	0.23	0.32	0.38**	0.49**	0.50**	0.52**
4. SPQ Ideas of reference				-	0.61**	0.41**	0.35**	0.36**	0.61**	0.41**	0.50**
5. SPQ Suspiciousness					-	0.43**	0.48**	0.58**	0.44**	0.31	0.40**
6. SPQ Social anxiety						-	0.36**	0.46**	0.32	0.14	0.19
7. SPQ Constricted affect							-	0.57 **	0.46**	0.19	0.35**
8. SPQ No close friends								-	0.39**	0.19	0.26
9. SPQ Odd speech									-	0.33	0.55**
10. SPQ Odd beliefs										-	0.52**
11. SPQ Odd behaviour											-

Note: ** indicates significance at $p < 0.001$;

Table 2.2. Correlations between RFQ dimensions and SPQ sub-scales

Abbreviations: RFQu, RFQ Uncertainty about Mental States; RFQc, RFQ Certainty about Mental States; SPQ, Schizotypal Personality Questionnaire

There were no concerns for multicollinearity between the independent variables. Multiple linear regression analyses indicated that when all SPQ subscales were entered together as predictors in the model, only the social anxiety ($\beta = 0.26$, $t = 2.37$, $p < 0.05$) and odd speech SPQ subscales ($\beta = 0.25$, $t = 2.02$, $p < 0.05$) significantly accounted for increased RFQu scores, controlling for the effects of age and YSR/ASR anxious/depressed scale scores (table 2.3.).

Furthermore, regression analyses showed that increases in SPQ odd speech significantly accounted for decreased RFQc scores ($\beta = -0.30$, $t = -2.31$, $p < 0.05$), controlling for the effects of age and YSR/ASR anxious/depressed scale scores (table 2.4.).

Dependent Variables	Model		Coefficients			Collinearity Statistics	
Control and Independent Variables	R^2	F	B	t	P	Tolerance	VIF
RFQu	0.29	3.37			<0.01		
Age			0.007	0.06	0.95	0.67	1.50
YSR/ASR Anxious/depressed			0.17	1.53	0.13	0.61	1.64
SPQ Social Anxiety			0.26	2.37	<0.05	0.66	1.52
SPQ Odd speech			0.25	2.02	<0.05	0.50	2.02
SPQ Constricted affect			0.06	0.55	0.58	0.61	1.65
SPQ Ideas of reference			0.03	0.19	0.85	0.36	2.81
SPQ Lack of close friends			-0.02	-0.14	0.89	0.46	2.20
SPQ Suspiciousness			0.08	0.58	0.56	0.38	2.63
SPQ Odd beliefs			-0.04	-0.29	0.77	0.56	1.80
SPQ Odd behaviour			0.03	0.25	0.80	0.48	2.09
SPQ Unusual perceptual experiences			-0.10	-0.79	0.43	0.44	2.29

Table 2.3. Linear multiple regression of SPQ on RFQu

Abbreviations: RFQu, RFQ Uncertainty about Mental States; SPQ, Schizotypal Personality Questionnaire; YSR/ASR, Youth and Adult Self Report

Dependent Variables	Model		Coefficients			Collinearity Statistics	
Control and Independent Variables							
	R^2	F	B	t	P	Tolerance	VIF
RFQc	0.20	2.08			<0.05		
Age			0.19	1.66	0.10	0.67	1.50
YSR/ASR Anxious/depressed			0.18	1.55	0.13	0.61	1.64
SPQ Social Anxiety			-0.07	-0.64	0.53	0.66	1.52
SPQ Odd speech			-0.30	-2.31	<0.05	0.50	2.02
SPQ Constricted affect			-0.24	-1.97	0.05	0.61	1.65
SPQ Ideas of reference			0.07	0.44	0.66	0.36	2.81
SPQ Lack of close friends			-0.02	-0.14	0.89	0.46	2.20
SPQ Suspiciousness			0.06	-0.40	0.69	0.38	2.63
SPQ Odd beliefs			-0.02	-0.19	0.85	0.56	1.80
SPQ Odd behaviour			0.93	0.69	0.49	0.48	2.09
SPQ Unusual perceptual experiences			-0.01	-0.06	0.95	0.44	2.29

Table 2.4. Linear multiple regression of SPQ on RFQc

Abbreviations: RFQu, RFQ Uncertainty about Mental States; SPQ, Schizotypal Personality Questionnaire; YSR/ASR, Youth and Adult Self Report

A series of Bayesian multi-model linear regressions with RFQ scales as dependent variables and SPQ subscales as independent variables were conducted to explore the results of the main regression analyses presented in tables 2.3. and 2.4.

First, Bayesian statistics confirmed the findings of the main regression analyses with RFQu as the dependent variable and SPQ subscales as independent variables, showing that the data increased the odds of including SPQ social anxiety ($BF_{inclusion} = 6.68$) and SPQ odd speech ($BF_{inclusion} = 3.20$) as predictors of RFQu. In contrast, after observing the data the inclusion probabilities of all other SPQ subscales as predictors of RFQu were reduced ($BF_{inclusion} < 0.30$).

In terms of RFQc, results of Bayesian analyses confirmed the findings of the main analyses by showing that the data increased the prior odds of including SPQ odd speech as a predictor of RFQc ($BF_{inclusion} = 1.78$). Interestingly however, the odds of including SPQ

constricted affect as a predictor of RFQc also showed a small increase ($BF_{inclusion}$ = 1.20). Given that results of the main regression analyses presented in table 2.3 showed a trend level association between SPQ constricted affect and RFQc ($p = 0.05$), it remains possible that with a larger sample this effect may have reached statistical significance. In contrast, after observing the data the inclusion probabilities of all other SPQ scales as predictors of RFQc were reduced ($BF_{inclusion} < 0.23$).

2.3.3 Effects of schizotypal traits and mentalizing on thought problems

Two multiple linear regressions were computed to examine the effects of (a) SPQ subscale scores and (b) RFQ scale scores, on the YSR/ASR thought problems subscale scores, controlling for the effects of age. Results showed that only SPQ social anxiety ($\beta = 0.17$, $t = 2.02$, $p < 0.05$) and SPQ odd speech ($\beta = 0.28$, $t = 2.86$, $p = 0.005$) were significantly associated to YSR/ASR thought problems. In terms of mentalizing, only RFQu was significantly associated with YSR/ASR thought problems ($\beta = 0.38$, $t = 3.92$, $p < 0.001$) (Table 2.5.).

Dependent Variables	Model		Coefficients			Collinearity Statistics	
Control and Independent Variables							
	R^2	F	B	t	P	Tolerance	VIF
YSR/ASR Thought Problems	0.56	11.72			<0.001		
Age			0.07	0.91	0.37	0.92	1.09
SPQ Social Anxiety			0.17	2.02	<0.05	0.66	1.52
SPQ Odd speech			0.28	2.86	<0.01	0.50	2.02
SPQ Constricted affect			-0.05	-0.60	0.55	0.62	1.60
SPQ Ideas of reference			0.04	0.33	0.75	0.36	2.80
SPQ Lack of close friends			0.12	1.15	0.26	0.46	2.17
SPQ Suspiciousness			0.09	0.85	0.40	0.38	2.61
SPQ Odd beliefs			0.06	0.71	0.48	0.59	1.71
SPQ Odd behaviour			0.12	1.19	0.24	0.48	2.09

SPQ Unusual perceptual experiences			0.17	1.61	0.11	0.44	2.28
YSR/ASR Thought Problems	0.19	8.06			<0.001		
Age			0.12	1.31	0.19	0.99	1.00
RFQu			0.38	3.92	<0.001	0.84	1.20
RFQc			-0.20	-0.98	0.33	0.83	1.20

Table 2.5. Effects of SPQ subscales and RFQ scales on YSR/ASR Thought Problems

Abbreviations: RFQu, RFQ Uncertainty about Mental States; RFQc, RFQ Certainty about Mental States; SPQ, Schizotypal Personality Questionnaire; YSR/ASR, Youth and Adult Self Report

To confirm the results of the regression analyses presented in table 2.5, Bayesian multi-model linear regressions were conducted to explore the effects of (a) SPQ subscale scores and (b) RFQ scale scores on YSR/ASR thought problems subscale scores, controlling for the effects of age. In terms of the effects of SPQ subscales on YSR/ASR thought problems, results of the Bayesian regression confirmed the findings of the main analyses, showing that the data increased the prior odds of including SPQ social anxiety ($BF_{inclusion} = 2.39$) and SPQ odd speech ($BF_{inclusion} = 45.76$) as predictors of YSR/ASR thought problems. In contrast, after observing the data the inclusion probabilities of all other SPQ scales as predictors of YSR/ASR thought problems were reduced ($BF_{inclusion} < 1.00$).

Similarly, Bayesian regression analyses with RFQ scales as independent variables and YSR/ASR thought problems as the dependent variable confirmed the results of the main analyses by showing that the prior odds of including RFQu as a predictor of YSR/ASR thought problems increased ($BF_{inclusion} = 492.74$), while the odds of including RFQc as a predictor decreased ($BF_{inclusion} = 0.34$).

2.3.4 Mediation analysis

According to the results of the regression analysis presented in Table 2.5., the social anxiety

and odd speech SPQ subscales, as well as RFQu significantly accounted for YSR/ASR thought problems scale scores. Thus, two mediation models were tested with YSR/ASR thought problems as the outcome variable. In the first model SPQ social anxiety was entered as the independent variable and RFQu was entered as the mediator. In the second model SPQ odd speech was entered as the independent variable and RFQu was entered as the mediator. Age was entered as a control variable in both models.

When controlling for age, SPQ social anxiety was significantly associated to RFQu (b = 0.47, 95% CI [0.23,0.70], p < 0.001) (path a in Fig. 2.2.) and RFQu was significantly associated to thought problems (b = 0.90, CI [0.25, 1.54], p < 0.01), (path b in Fig. 2.2). There was a significant indirect effect of SPQ social anxiety on YSR/ASR thought problems through RFQu (b = 0.42, 95% confidence interval [0.1370, 0.9145], p < 0.05), indicative of a mediation effect. The total effect of SPQ social anxiety on YSR/ASR thought problems (b = 1.68, 95% confidence interval [1.03, 2.33], p < 0.0001) (path c in Fig. 2.2.) was reduced by 25% when RFQu was accounted for (b = 1.26, 95% confidence interval [0.72, 1.80], p < 0.0001) (path c' in Fig. 2.2.), indicating partial mediation.
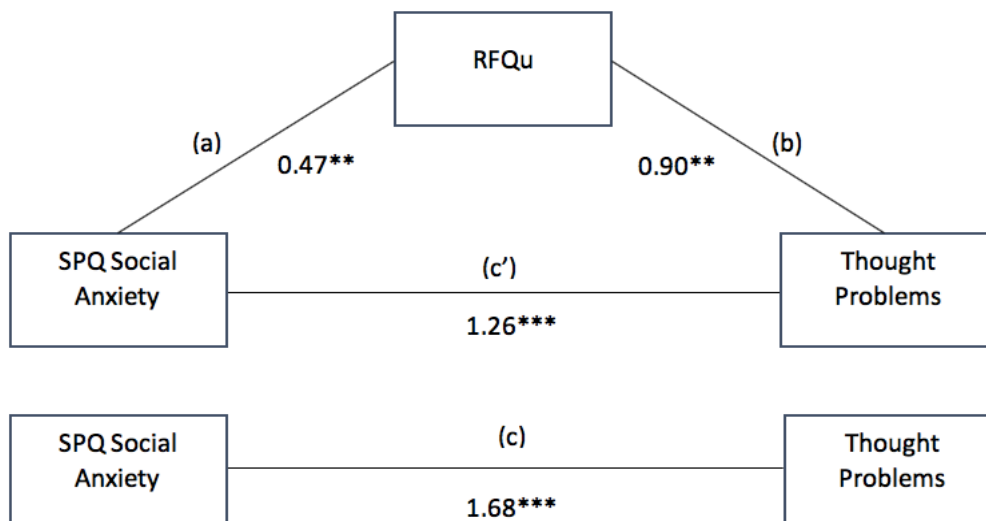
Figure 2.2. Mediational model with SPQ Social Anxiety as independent variable, RFQu as mediator and thought problems as dependent variable
Abbreviations: SPQ, Schizotypal Personality Questionnaire; RFQu, Reflective Functioning Questionnaire uncertainty about mental states scale

Results further indicate that when controlling for age, SPQ odd speech was significantly associated to RFQu ($b = 0.40$, 95% CI [0.23,0.58], $p < 0.0001$) (path a in Fig. 2.3), and RFQu was significantly associated to YSR/ASR thought problems ($b = 0.68$, CI [0.12, 1.23], $p < 0.05$) (path b in Fig. 2.3). There was a significant indirect effect of SPQ odd speech on YSR/ASR thought problems through RFQu ($b = 0.27$, 95% confidence interval [0.05, 0.60], $p < 0.05$), indicative of a mediation effect. The total effect of SPQ odd speech on YSR/ASR thought problems ($b = 1.88$, 95% confidence interval [1.22, 2.54], $p < 0.0001$) (path c in Fig. 2.3) was reduced by 14% when RFQu was accounted for ($b = 1.61$, 95% confidence interval [0.99, 2.23], $p < 0.0001$) (path c' in Fig. 2.3), indicating partial mediation.
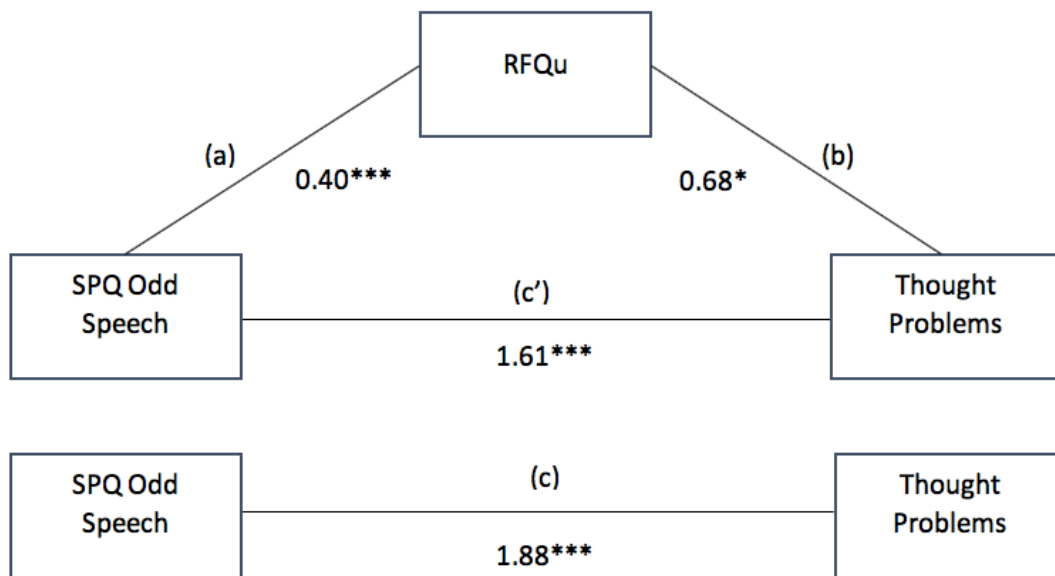
Figure 2.3. Mediational model with SPQ Odd Speech as independent variable, RFQu as mediator and thought problems as dependent variable

Abbreviations: SPQ, Schizotypal Personality Questionnaire; RFQu, Reflective Functioning Questionnaire uncertainty about mental states scale

2.4 Discussion

The first aim of the current study was to identify schizotypal features that contribute to mentalizing difficulties in a sample of community youths. Furthermore, the current study

examined the effects of schizotypal features and mentalizing on self-reported thought problems. Finally, the mediating role of mentalizing on the relationship between schizotypal features and thought problems was tested.

2.4.1 Associations between schizotypal personality features and self-reported mentalizing

In accordance with the hypothesis, the present findings show that interpersonal schizotypal manifestations pertaining to social anxiety are associated to high mentalizing uncertainty. This resonates with data suggesting that mentalizing deficits pertaining to the understanding of others' mental states as assessed by ToM tasks, are linked with the level of interpersonal withdrawal in patient samples (Frith, 2014; Sergi et al., 2007; Weijers et al., 2018). Current results are also consistent with recent findings demonstrating links between diminished mentalizing capacities in adolescence and retrospective reports of social anxiety in childhood (Ballespí et al., 2018). Young people who withdraw from interpersonal contact in the context of social anxiety, may encounter less social interaction opportunities, essential for developing a capacity to understand mental states (Banerjee & Henderson, 2001). Alternatively, or in conjunction, difficulties in understanding the mental states of oneself and others may interfere with young people's ability to function within social situations, leading to experiences of increasing anxiety in the context of interpersonal contact. Indeed, according to Banerjee and Henderson (2001), the relationship between social anxiety and mentalizing dysfunction appears to be bidirectional (one influences the other across development). For those developing clinical psychosis, poor mentalizing appears to further impede interpersonal communication, increase the tendency to withdraw from social situations, and reduce treatment engagement (Debbané et al., 2016).

A novel finding of the present study is that schizotypal personality features of odd speech are associated with high mentalizing uncertainty and reduced mentalizing certainty in adolescence. As in the case of social anxiety, speech that is vague, concrete, or tangential, impedes the effective communication of thoughts and feelings within interpersonal interactions, therefore potentially disrupting the developmental elaboration of complex meta-representations of self-experience. Given that deficits in the communicative use of speech are prominent in schizophrenia and reflect a genetic vulnerability for the illness (Docherty & Gottesman, 2000), the association between schizotypal manifestations of odd speech and mentalizing difficulties in adolescence may represent a distal marker of psychosis vulnerability.

Contrary to the present findings, previous studies examining the relationship between mentalizing abilities and schizotypal traits using ToM-tasks have reported associations with cognitive-perceptual manifestations (Pickup, 2006). This could be attributed to the different methods employed to assess mentalizing (ToM vs. RFQ). Task-based measures of ToM, mainly capture the ability to form inferences about other peoples' cognitions, and it has been suggested that contrary to clinical psychosis, poor ToM in schizotypy may be primarily related to hallucination and delusion-like phenomena (Pickup, 2006). Conversely, the RFQ is a self-report measure designed to also assess self-oriented and affect-based facets of mentalizing, as well as the ability to make mentalizing inferences of the self in relation to close others, thus may be more sensitive in capturing associations with interpersonal features of schizotypy. Importantly, the only study to date that has examined the relationship between ToM and schizotypy in community adolescents also found an association with the cognitive-perceptual dimension of schizotypy (Barragan et al., 2011). Thus, differences between the current findings and those reported in adult ToM studies cannot be readily attributed to the inclusion of an

adolescent sample. Future studies should examine the associations between RFQ-measured mentalizing and schizotypy in adult samples.

2.4.2 Effects of schizotypal personality features and self-reported mentalizing difficulties on thought problems in adolescence

Current findings suggest that schizotypal trait features pertaining to social anxiety and odd speech are associated to the level of thought problems in a sample of community adolescents. The association between social anxiety and thought problems in the current sample is in line with previous research indicating that social anxiety relates to the severity of psychotic symptoms in CHR adolescents and young adults (Pontillo et al., 2017). Data from the present study are also in line with previous studies reporting that disorganized speech relates to thought problems in both chronic psychosis and non-clinical samples (Sommer et al., 2010).

Interestingly, results indicate that mentalizing uncertainty mediated the effects of schizotypal features on the level of thought problems in the sample. These were only partial mediations, suggesting that other interpersonal, cognitive and psychological factors may also shape the relationship between schizotypal traits and thought problems in adolescence. Nevertheless, current findings lend support to a model in which mentalizing difficulties contribute to the relation between schizotypal traits and thought problems that lie in a continuum with clinically-relevant psychotic symptoms (Debbané & Barrantes-Vidal, 2014; Debbané, Salaminios et al., 2016).

First, findings suggest that mentalizing uncertainty may represent a potential pathway through which trait features of social anxiety relate to thought problems in adolescence. Social anxiety has been consistently linked to the development and maintenance of psychotic symptoms (Rietdijk et al., 2013), including paranoid/persecutory ideation (Michail and Birchwood, 2009), however the underlying mechanisms remain incompletely understood. One

possibility is that difficulties in understanding the mental states of others in the context of social anxiety may lead individuals to overattribute other people's intentions (Ballespí et al., 2019). Indeed, the tendency to overattribute mental states in others, also known as *"hyper-mentalizing"* (Sharp et al., 2013), has been associated with both paranoid beliefs and hallucination-like phenomena in community samples (Clemmensen et al., 2014) and in adults suffering with schizophrenia (Ciaramidaro et al., 2015). Although the use of the RFQ in the current non-clinical sample did not permit for the direct assessment of hyper-mentalizing misattributions, previous studies suggest that difficulties in understanding mental states may trigger the over-attribution of intentions in others, leading to clinically-relevant manifestations (Langdon and Rock, 2008).

Current findings also suggest that uncertainty in mental states partially mediated the relationship between schizotypal features of odd speech and thought problems in our sample. It must be noted however that the indirect effects of the mediational model were small. Indeed, it has been reported that the relationship between disorganised speech and clinically-relevant thought problems, such as hallucinatory phenomena, may primarily reflect the outcome of a shared neurobiological basis between the two (Sommer et al., 2010). Importantly though, disorganisation features of schizotypy (including odd speech) in adolescence have been longitudinally associated with the developmental trajectory of clinically-relevant cognitive-perceptual manifestations (Dominguez et al., 2010). The present findings add to this literature, suggesting that the effect of disorganised speech on psychosis-risk may partly depend upon the level of mentalizing impairment.

Mentalizing difficulties are transdiagnostic and have been implicated in the development and phenomenology of various psychiatric illnesses, such as BPD (Bateman & Fonagy, 2010); depression (Taubner et al., 2011) and functional somatic disorders (Luyten et al., 2012). Regarding psychotic disorders, results of the current study tentatively suggest that

mentalizing dysfunction may represent a potential pathway through which schizotypal personality traits affect the level of thought problems in adolescence. From an early prevention standpoint, mentalizing difficulties may constitute important psychotherapeutic targets to sustain resilience against psychosis-risk during adolescent development.

2.4.3 Limitations and conclusions

The results of the current study should be interpreted in light of certain limitations. First, the data were derived from a relatively small sample and no prior power analyses were conducted to determine the sample size needed to detect significant effects. Thus it remains possible that further associations could have emerged with a larger sample. It must be noted however that to further explore whether the results of the main analyses represented true effects as opposed to non-conclusive findings and to reduce the probability of type II errors, Bayesian statistics were obtained and overall these provided support in favour of the study's findings. Nonetheless, future studies should test the results of the current study with larger samples. Second, the analyses performed were cross-sectional and longitudinal investigations are needed to establish the causal relevance of schizotypy on the developmental cascade of mentalizing. Third, mentalizing represents a multidimensional construct reflecting the capacity to make implicit and explicit inferences about one's own and other people's cognitions and affects (Fonagy and Luyten, 2009). However, the RFQ is a screening measure, not designed to capture different dimensions of mentalizing. Therefore, the assessment of specific mentalizing dimensions and their relation to schizotypy or thought problems in the current sample was not possible. In addition, given that the study was based on self-report measures, the data may have been influenced by the accuracy with which young people evaluated their mentalizing abilities. Indeed, difficulties in reflecting on one's own mental states may have led to biased responses in the RFQ (Fonagy et al., 2016; Shaw et al., 2019). Finally, the study could have benefited by

including measures of interpersonal factors known to influence mentalizing processes and psychosis vulnerability, such as childhood trauma (Weijers et al., 2018) and insecure attachment (Debbané et al., 2016).

Despite these limitations, this study is the first to examine the associations between mentalizing difficulties and schizotypal features during adolescence, as well as their effect on thought problems, using an easy-to-administer measure of mentalizing. The findings underlie the importance of encompassing evaluations of mentalizing when clinically assessing psychosis-risk in youths who present schizotypal features. Mentalization-based treatment (Brent & Fonagy, 2014; Debbané et al., 2016) may be applied preventatively to sustain mentalizing functioning and attenuate psychosis-risk in adolescents who present schizotypal manifestations, prior to the development of clinical symptoms.

Chapter 3

Self-Monitoring for Speech and its Links to Age, Cognitive Effort, Schizotypal Trait Expression and Impulsivity During Adolescence[2]

3.1. Introduction

According to contemporary conceptualizations, clinical manifestations of psychotic illnesses, particularly those that entail cognitive-perceptual aberrations, are associated with confusions in the identification and monitoring of self-generated information. For instance, auditory hallucinations have been linked with difficulties in recognising and monitoring the origin one's own inner speech (Allen et al., 2007), while delusions of alien control have been attributed to difficulties in monitoring the initiation and sensory consequences of self-produced actions (Blakemore et al., 2003). As discussed in chapter 1.6., a key cognitive process associated with the monitoring of self-generated material is the ability to discriminate between different sources of memory information, also known as *source monitoring* (Johnson et al., 1993).

To date, two types of source-monitoring processes have consistently been examined in people suffering with psychotic disorders: (1) *reality-monitoring*, which involves the capacity to distinguish between self-generated and externally-derived information; and (2) *self-monitoring* (or internal source-monitoring), which involves the capacity to discriminate between two internal sources of information (what one imagined doing or saying from what one did or said). Although both processes characterise the capacity to discriminate between internal and external information (thus self-monitoring may in itself be considered as an aspect of reality-monitoring), they can also be conceptually distinguished according to whether a particular investigation takes as its focus (a) the self versus an external source as origin of

---

information (i.e. what one said aloud or thought from what one heard), or (b) the covert/non-public origin of mental events (e.g. imagination) versus events that have adopted an overt/public quality (Johnson et al., 1993). According to the *source-monitoring framework* (Johnson et al., 1993), both reality- and self-monitoring judgments about the source of memory information are based on evaluations pertaining to the characteristics of the memory events themselves. These primarily include the sensory characteristics (visual/auditory) present and the record of cognitive operations engaged during the original encoding of an event, as well as the activation of these during memory retrieval. For instance, studies suggest that memories for externally-perceived events contain more sensory information compared to memories for imagined events and the capacity to discriminate between the two increases with the amount and clarity of such detail contained in the former (Dewhurst, 1999; Johnson, 1988). Conversely, because imagination is less automatic than perception, memories for imagined events contain more records about the cognitive operations that led to their creation compared to those for real events (Sussman, 2001). Indeed, increased cognitive effort during memory encoding for imagined events has been shown to support accurate source-monitoring discriminations (Finke et al., 1988).

Regarding impairments in source memory, the source-monitoring framework postulates that because contextual characteristics and cognitive operations are used as cues that determine the origin of memory events, source confusions are likely to occur under conditions that alter the quantity and quality of such information during encoding or retrieval (Johnson et al., 1993). For example, studies indicate that conditions that increase the sensory information encoded for imagined events typically lead to their misattribution as externally-perceived and real (Johnson, 1988). Similarly, increases in the cognitive elaboration required during the encoding of externally-perceived events commonly lead to their misattribution as self-generated and imagined (Johnson, 2006). Although source memory is an imperfect process and

transient source confusions are relatively common, impairments in source-monitoring have also been linked with the phenomenology of psychotic illnesses.

Reality-monitoring research in psychosis has shown that patients suffering with schizophrenia and individuals at clinical high-risk for the illness (CHR) display a tendency towards misattributing the source of self-generated items to external agents (Bentall et al., 1991; Brébion et al., 2000; Johns et al., 2010; Johns et al., 2006). This pattern of externalizing reality-monitoring misattributions is particularly pronounced among patients reporting hallucinations (Bentall et al., 1991; Brébion et al., 2000) and non-clinical adults experiencing hallucination-like phenomena (Larøi et al., 2004). Interestingly, increased cognitive effort appears to exacerbate the rate of externalizing misattributions in these samples (Bentall et al., 1991; Larøi et al., 2004). This suggests that unlike healthy controls, schizophrenia sufferers and high-risk subjects fail to utilize cognitive effort as a cue to monitor the source of self-generated information (Larøi et al., 2004).

Given that psychotic experiences entail pervasive difficulties in discriminating between imagination and reality, studies have also investigated self-monitoring processes in schizophrenia. These have addressed whether participants can distinguish between two internal sources of information (what one did from what one thought or imagined doing). For example, Keefe et al. (1999) found that together with difficulties in discriminating between self-generated and experimenter-produced words, patients with schizophrenia also displayed confusions in memory between words they had read aloud and words they read silently. Most specifically, in line with reality-monitoring research, self-monitoring studies suggest that patients with schizophrenia and CHR individuals demonstrate externalizing misattributions towards recalling imagined stimuli as overtly enacted (Docherty, 2012; Franck et al., 2000; Gawęda et al., 2012; Henquet et al., 2005).

According to Docherty (2012), source discriminations for "internal" events are never purely internal. This is because once internal events are enacted, they also become part of the perceptually-based external world. For instance, contrary to silent speech, overt speech entails the generation of motor actions (i.e. vocalizations), along with the perception of the sensory feedback produced by these actions (i.e. listening to the spoken material). In the same vein, Johnson et al. (1993) indicate that discriminating in memory between covert/imagined mental events (i.e. silent speech) and events that have adopted a public quality (i.e. overt speech) may in itself constitute an instance of reality-monitoring. Therefore, the overlap between various types of source discrimination problems in psychosis may signify the presence of a more global deficit in the processing of internal and external information (Docherty, 2012). Indeed, according to the source-monitoring framework, source misattributions occur due to alterations in the encoding of sensory and cognitive information (Johnson et al., 1993). This is also in line with the *forward model of motor control* (Frith et al., 2000), which postulates that source confusions result from failures to match the intended or predicted outcomes of self-generated actions to their externally-perceived sensory consequences. This can occur either due to disruptions in generating predictions about the outcome of actions, or due to impairments in the processing of the sensory feedback related to these actions. In both cases, because of the high level of sensory discrepancy between cognitive prediction and sensory feedback, self-generated events are experienced as external to the self. Taken together, the source-monitoring framework and the forward model suggest that the self-monitoring confusions observed in psychotic illnesses are sustained by impairments in the monitoring of sensory and cognitive information.

Another line of research indicates that self-monitoring impairments akin to those identified in people suffering with psychotic disorders are also observable, albeit at an attenuated level, among non-clinical individuals in the context schizotypal trait expression,

prior to development of clinical symptoms (Peters et al., 2007). Schizotypal manifestations represent distal risk factors for schizophrenia (Debbané et al., 2014) and most psychometric analyses examining their factorial structure identify three dimensions: the cognitive-perceptual (positive schizotypy: hallucination, delusion-like phenomena), interpersonal (negative schizotypy: social anxiety, constricted affect), and disorganization dimensions (odd behaviours, odd speech) (Fonseca-Pedrero et al., 2018; Raine, 2006). Only a limited number of studies to date have analysed the associations between schizotypal dimensions and self-monitoring in non-clinical individuals. Peters et al. (2007) examined the relationship between action self-monitoring (distinguishing between performed physical actions versus imagined actions) and positive schizotypy in healthy adults ($M_{age} = 21.41$). They found that participants reporting high positive schizotypy displayed more externalizing misattributions for actions (confusing imagined actions as overtly enacted), compared to those with lower scores. More recently, Humpston et al. (2017) found that cognitive-perceptual schizotypal manifestations in community adults ($M_{age} = 22.30$) were positively associated with the total rate of action self-monitoring errors. Although suggestive of a relationship between self-monitoring and premorbid trait signs of psychosis, these studies involve adult participants either within or beyond the critical period of illness onset. From an early intervention standpoint, important insights can be gained by examining self-monitoring processes in the context of schizotypal trait expression during earlier developmental stages.

Developmental research suggests that self-monitoring abilities improve between early childhood and young adulthood (Markham, 1991; Folley et al., 1983). Interestingly however, little is known at present about the developmental trajectories of self-monitoring processes during adolescence. This is important as adolescence also represents a key period during which schizotypal manifestations become clinically relevant (Gooding et al., 2005). Debbané et al. (2010) reported that adolescents ($M_{age} = 13.97$) prone to express positive schizotypal

manifestations due to 22q11.2 deletion syndrome, a neurogenetic condition conferring high-risk for schizophrenia, exhibited more externalizing and internalizing self-monitoring speech misattributions compared to an adolescent control group. Similarly to clinical adult samples, increased cognitive effort exacerbated the rate of self-monitoring errors in adolescents with 22q11.2 deletion syndrome (Debbané et al., 2010). Despite these findings, no study to date has directly examined the relationship between schizotypal dimensions and disruptions in self-monitoring among typically developing adolescents. Research on adolescent schizotypy, which precedes the emergence of clinical manifestations, can increase our understanding regarding the earliest phases of self-monitoring impairments in schizophrenia, with implications for early prevention treatment.

Impulsivity represents another personality feature associated with the emergence of clinical psychopathology in adolescence and young adulthood (Moeller et al., 2001). Impulsivity constitutes a multidimensional trait reflecting the tendency to act rashly without regard to negative consequences (Moeller et al., 2001). According to Whiteside and Lynam (2001) impulsivity is comprised of four dimensions: urgency (tendency to act rashly under positive or negative affect); lack of premeditation (tendency to not think about consequences of actions); lack of perseverance (inability to remain focused); and sensation seeking (tendency to pursue new and exciting activities). Evidence suggest that increased levels across impulsivity dimensions lie at the centre of multiple problematic behaviours arising in adolescence, such as antisocial behaviour, substance abuse, risk-taking and delinquency (Moeller et al., 2001). Although most psychological problems associated with impulsivity also entail cognitive difficulties (Müller et al., 2015), little is known about the relationship between self-monitoring and impulsivity. Recent data suggest that incarcerated adolescents that commonly report higher levels of impulsivity compared to community adolescents, exhibit increased rates of externalizing and decreased rates of internalizing self-monitoring

misattributions (Morosan et al., 2018). Interestingly however, the only study that has examined the relationship between impulsivity and schizotypy in adolescence has reported inverse associations between the two (Badoud et al., 2015). Thus, self-monitoring may represent a transdiagnostic cognitive mechanism associated to different personality traits involved in the emergence of both psychotic and non-psychotic manifestations (Ferchiou et al., 2010). At present, studies encompassing evaluations of impulsivity when examining self-monitoring in adolescence are lacking.

The present study seeks to examine self-monitoring in community adolescents using a task that assesses the capacity to discriminate between self-generated overt and silent speech under different levels of cognitive effort (Debbané et al., 2010). The main aims of the current study are to assess the effects of age and cognitive effort on self-monitoring for speech in adolescence, as well as its associations with personality dimensions pertaining to schizotypy and impulsivity. Furthermore, to explore whether recognition and source memory draw on similar developmental and heuristic processes, particularly as it pertains to verbally-presented stimuli, the study will also examine the effects of age and cognitive effort on recognition performance.

Given that verbal memory commonly improves during adolescence (Murre et al., 2013) it is expected that older adolescents will exhibit better recognition for speech than younger adolescents. Furthermore, because overt reading involves the encoding of sensory-motor signals, which can be used to distinguish previously presented from new stimuli, it is hypothesized that participants will show better recognition for overtly-read than silently-read items. Regarding self-monitoring, previous research indicates that young children perform worse than older children and adults in self-monitoring for speech (Foley et al., 1983; Sussman, 2001). Thus, it is hypothesised that younger adolescents will exhibit more self-monitoring errors compared to older adolescents. In accordance with the source-monitoring framework, it

is also expected that the current sample will demonstrate increased rates of internalizing misattributions under high cognitive effort, which generates more cognitive operations during encoding. In terms of the associations with schizotypal traits, following previous action self-monitoring research in adults (Humpston et al., 2017; Peters et al., 2007), it is hypothesised that the rate of externalizing misattributions for speech will be positively associated with the cognitive-perceptual dimension of schizotypy. Given that the majority of self-monitoring studies have focused on cognitive-perceptual manifestations and no study to date has examined self-monitoring for speech and it's links to schizotypal trait expression in community adolescents, the present study also seeks to explore whether higher scores across different schizotypy dimensions, including the interpersonal and disorganization dimensions, will be associated with increases in the total rate of self-monitoring misattributions (used in the current study as a measure of general self-monitoring performance). Finally, on the basis of previous work involving incarcerated youths (Morosan et al., 2018) it is expected that impulsivity dimensions will show positive associations with externalizing misattributions and negative associations with internalizing misattributions.

3.2. Methods

3.2.1. Participants and procedure

A convenience sample of 139 community adolescents was recruited through written advertisements in public schools in the city of Geneva, Switzerland.. No prior power analysis was performed to determine the size of the sample. Inclusion criteria were age (12-18 years) and fluency in French. Participants were screened for cognitive impairment. Those with scores below a standard score of 7 in the Block Design subtest of the Wechsler Intelligence Scale (Wechsler, 1955) were excluded from the analyses (n = 17). To account for insufficient attention to the self-monitoring task, participants scoring three standard deviations below the

mean *d'* score in the recognition test were also excluded (n = 1). None of the participants scored three standard deviations above the mean *d'* score. The final sample consisted of 121 adolescents (56 females, 65 males, M_{age}= 15.23). None of the participants suffered from past/present psychiatric, or neurological/neurogenetic disorders. Furthermore, none of the participants of the current study participated in study 1 the current thesis. Written informed consent was obtained from all participants and legal guardians. Ethical approval for the study was granted by the Swiss Ethics Committee on research involving humans (number: 2018-00251)

3.2.2. Measures

3.2.2.1. Self-monitoring task

Self-monitoring was investigated using a speech-monitoring task (Debbané et al., 2010) consisting of two parts: a reading procedure, followed by a recognition and self-monitoring test. In the first phase participants were required to read, either aloud or silently, a series of common words (low cognitive effort) or non-words (high cognitive effort) presented on a computer screen. The inclusion of non-word items in the high-cognitive effort condition was based on neuroimaging research on language processing, suggesting that the reading of non-words, which contain unfamiliar phonological associations and no immediate semantic associations, entails more effortful processing and stronger neural activation than that of words (Price & Devlin, 2011; Taylor et al., 2013).

During the procedure, participants were seated 60 cm away from a computer screen. The presentation of stimuli was administered by E-prime software and all items were presented in black letters (Courier New font; size 18) on a white background. A total of 72 words with comparable frequency and usage in French language were selected for the task. Half of these words were used as the word items and the other half were used as the source for the non-word

items. Item letter length (consonant = c; vowel = v) was either four (c-v-c-v), five (c-v-c-v-c), or six letters (c-v-c-v-c-v). Non-word items were produced according to the following procedure: for c-v-c-v items the syllables were inversed; for c-v-c-v-c items the last two consonants were inversed; and for c-v-c-v-c-v items the first two syllables were inversed (see Appendix 4)

The first part of the task was presented to participants as a reading and pronunciation exercise. Subjects were asked to pay particular attention to their pronunciation, even during silent reading and were not informed that a recognition and self-monitoring test would follow. During the reading procedure, six blocks (three silent, three aloud) of eight items (four words, four non-words) were randomly presented. Each condition contained 12 items, for a total of 48 items (12 word/overt items, 12 non-word/overt items, 12 words/silent items and 12 non-word/silent items). After a 15-minute filler task, a recognition sheet was handed out, containing 72 items (48 original items,12 new word and 12 new non-word items). In the test phase, participants had to indicate if each item from the list appeared in the first part of the task (yes/no - recognition test), and attribute recognized items to a reading condition, silent or aloud (self-monitoring test). The task yields two types of self-monitoring errors: externalizing errors (attributing silently-read items as overtly-read), and internalizing errors (attributing overtly-read items as silently-read); for two item types involving different cognitive effort levels: high cognitive effort (non-words) and low cognitive effort (words).

For the recognition test, signal detection theory (Stanislaw & Todorov, 1999) was used to assess the sensitivity for each type of items (word and non-words) for both reading conditions (aloud and silently). To estimate d' scores the z score corresponding to false alarms was subtracted from the z score corresponding to hits.

The percentage of externalizing misattributions was calculated by dividing the total score for items read silently but identified as read aloud in the monitoring test, out of the total

number of correctly recognized silent items. The percentage of internalizing misattributions was calculated by dividing the total score of items read overtly but identified as read silently in the monitoring test, out of the total number of correctly recognized overt items. Percentages of externalizing and internalizing errors were calculated overall, as well as for each item type. Finally, the total percentage of self-monitoring errors was calculated as the sum of internalizing and externalizing errors for all items.

3.2.2.2. Self-report measures

The Schizotypal Personality Questionnaire (SPQ-French version, Dumas et al., 2000) measures schizotypal traits and their dimensions, yielding three factor scores: cognitive–perceptual (unusual perceptual experiences, suspiciousness, ideas of reference); interpersonal (social anxiety, constricted affect, lack of close friends); and disorganization (odd speech, odd behaviour). The French version of the SPQ has shown good reliability (Cronbach's alpha = 0.91) (Dumas et al., 2000) and has been validated for use with francophone adolescents (Badoud et al., 2011). The SPQ in the present study also showed good reliability (Cronbach's alpha = 0.86).

The UPPS Impulsive Behaviour Scale - short version (UPPS Billieux et al., 2012) measures four personality factors associated with impulsive behaviour: lack of premeditation (tendency to act without thinking); lack of perseverance (inability to remain focused in a task); urgency (tendency to act rashly under negative emotions); and sensation seeking (tendency to seek out novel and thrilling experiences). The UPPS has been validated for French-speaking samples, showing good reliability (Cronbach's alpha = 0.70-0.84) (Billieux et al., 2012).

3.2.3. Statistical analyses

For the self-monitoring task, a 2×2×3 mixed analysis of variance (ANOVA) was conducted on the d' recognition scores with two within subject factors (item type: word vs. non-word; and reading type: aloud vs. silently), and age-group as the between factor [12-14 years of age (early adolescents, n = 36), 14-16 years of age (middle adolescents, n = 39), 16-18 years of age (late adolescents, n = 46)]. A 2×2×3 mixed ANOVA was also conducted on self-monitoring error scores with two within subject factors: (monitoring error type: externalizing vs. internalizing; and item type: word vs. non-word), and age-group as the between factor [12-14 years of age (early adolescents), 14-16 years of age (middle adolescents), 16-18 years of age (late adolescents)]. Because male and female participants did not differ in self-monitoring scores, gender was not entered as a covariate in these analyses.

Because all SPQ and UPPS dimension scores violated the assumption of normality (Kolmogorov-Smirnov test, $p < 0.05$), Spearman's correlation coefficients were calculated to estimate their relationships with self-monitoring errors. To account for multiple comparisons and reduce the probability of type I errors Bonferroni adjusted significance levels were calculated.

Given that no prior power calculations were performed to determine sample size and in order to explore whether the results of the Spearman's correlations conducted represented true effects as opposed to non-conclusive findings JASP version 0.9.2 was used to obtain simple Bayesian statistics of the correlation analyses run (Nuzzo, 2017). In contrast to standard hypothesis testing based on P values, Bayesian statistics offer a direct measure of the strength of the evidence both in favour and against the study hypotheses that can support the interpretation of results. Specifically, Bayesian correlations provide a Bayes factor (BF_{10}) that represents the likelihood of the data under the null hypothesis (i.e. no correlation) divided by the likelihood of the data under the alternative (i.e. correlation present), so that BF_{10} values greater than 1 signal more evidence in favour of the alternative hypothesis. Furthermore,

Bayesian correlations provide another Bayes factor (BF_{01}) which represents $1/BF_{10}$, so that BF_{01} values greater than 1 signal more evidence in favour of the null hypothesis. Given that both BF_{10} and BF_{01} contain the same information, for ease of interpretation the results section will only report the Bayes factor that is greater than 1.

3.3. Results

3.3.1. Descriptive results

Table 3.1. presents the descriptive results for the variables included in the analyses. Kurtosis and skewness of the distributions for the self-monitoring task, SPQ and UPPS scales were all within the acceptable range (i.e. all values between -2.00 and 2.00) and no outliers were identified. Furthermore, analyses of the SPQ scale distributions showed that the proportion of participants reporting higher schizotypal trait scores (1.5 standard deviations above the normed means) was comparable across the different SPQ scales (9.92% for the interpersonal [n = 12]; 10.74% for the cognitive-perceptual [n = 13] and 8.26% for the disorganization [n = 10] SPQ dimensions. Similarly, the proportion of participants that reported impulsivity scores 1.5 standard deviations above the normed means was comparable across the different UPPS scales (9.09% for the lack of perseverance [n = 11]; 6.61% for the lack of premeditation [n = 8]; 8.26% for the urgency [n = 10] and 9.09% for the sensation seeking [n = 11] UPPS dimensions).

	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
Recognition Test						
Correct recognition/Hits raw score	29.37	5.69	13.00	44.00	-0.24	0.18
Correct recognition/Hits (%)	61.19	11.86	27.00	92.00	-0.24	0.18
False alarms raw score	4.59	3.00	0.00	15.00	0.90	0.96
False alarms (%)	19.21	12.41	0.00	63.00	0.90	0.96
d-prime read words	2.05	0.65	0.71	3.46	0.04	-0.68
d-prime silent words	1.33	0.53	0.24	1.33	-0.17	-0.85
d-prime read non-words	1.79	0.64	0.00	3.46	-0.35	0.35
d-prime silent non-words	1.47	0.61	0.21	3.11	0.12	-0.33
Self-monitoring Test						
Self-monitoring errors Total (%)	32.39	11.35	0.00	56.88	-0.21	-0.32
Internalizing errors Total (%)	34.83	14.48	0.00	68.18	-0.07	-0.55
Externalizing errors Total (%)	29.95	15.54	0.00	70.00	0.07	-0.41
Internalizing errors for words (%)	30.28	19.30	0.00	30.28	0.38	-0.02
Internalizing errors for non-words (%)	39.40	21.40	0.00	100.00	0.61	0.31
Externalizing errors for words (%)	32.24	22.90	0.00	100.00	0.61	0.48
Externalizing errors for non-words (%)	27.66	20.24	0.00	83.33	0.44	-0.28
Schizotypal Traits (SPQ)						
SPQ Cognitive-perceptual dimension	8.58	6.83	0.00	33.00	1.05	0.67
SPQ Interpersonal dimension	6.63	4.54	0.00	20.00	0.53	-0.35
SPQ Disorganization dimension	5.45	3.93	0.00	14.00	0.34	-1.02
Impulsivity (UPPS)						
UPPS Lack of Perseverance	2.04	0.60	1.00	4.00	0.45	0.33
UPPS Lack of Premeditation	2.17	0.61	1.00	4.00	0.54	0.56
UPPS Urgency	2.46	0.68	1.00	4.00	0.14	0.06
UPPS Sensation Seeking	2.69	0.69	1.00	4.00	-0.17	-0.67

Table 3.1. Means and standard deviations for each variable in the sample
Note: SPQ, Schizotypal Personality Questionnaire; UPPS, Impulsive Behaviour Scale

3.3.2. Self-monitoring task results

3.3.2.1. Recognition performance

Table 3.2. presents the age-group means and standard deviations for the *d'* scores, for each type of item in both reading conditions.

	Early Adolescents (n = 36 , age = 13.04)		Middle Adolescents (n = 39 , age = 15.13)		Late adolescents (n = 46 , age = 17.02)	
Recognition Test	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Correct recognition/Hits raw score	27.69	5.37	28.85	5.05	31.13	6.06
Correct recognition/Hits (%)	57.70	11.19	60.10	10.52	64.86	12.64
False alarms raw score	5.11	3.19	3.92	2.64	4.76	3.10
False alarms (%)	21.30	13.28	16.35	11.00	19.84	12.91
d-prime Total	1.51	0.29	1.68	0.30	1.76	0.36
d-prime read words	1.79	0.67	2.10	0.56	2.20	0.64
d-prime silent words	1.35	0.53	1.28	0.56	1.35	0.51
d-prime read non-words	1.50	0.63	1.80	0.65	2.02	0.55
d-prime silent non-words	1.40	0.56	1.53	0.60	1.47	0.66

Table 3.2. Age-group means and standard deviations for Hits, false alarm and d-prime scores

Results of the mixed ANOVA conducted on the *d'* recognition scores, with age-group as a between factor, reveal a main effect of age-group ($F_{(2, 118)} = 6.23$, $p = 0.003$). Post-hoc pairwise comparisons suggest that independently of reading condition and item type, the early adolescent group displayed significantly less accurate recognition scores compared to middle ($p = 0.02$) and late adolescents ($p = 0.001$). In addition, a significant interaction-effect was found between reading type and age-group ($F_{(2, 118)} = 5.34$, $p = 0.01$), suggesting that independently of item type, early adolescents less accurately recognized overtly-read items compared to middle ($p = 0.004$) and late adolescents ($p < 0.001$).

The mixed ANOVA conducted on d' scores also revealed a significant main effect of reading type ($F_{(1, 118)} = 85.42$, $p < 0.001$), suggesting that independently of age-group and item type, silently read items are less accurately recognized. Furthermore, results showed a

significant interaction effect between reading type and item type (F (1, 118) = 20.30, $p <$ 0.001), suggesting that independently of age-group, overtly-read words were significantly more accurately recognized than overtly-read non-words (F (1, 118) = 11.70, p = 0.001), while silently-read non-words were significantly more accurately recognized than silently-read words (F (1, 118) = 4.07, p = 0.046) (Figure 3.1.).
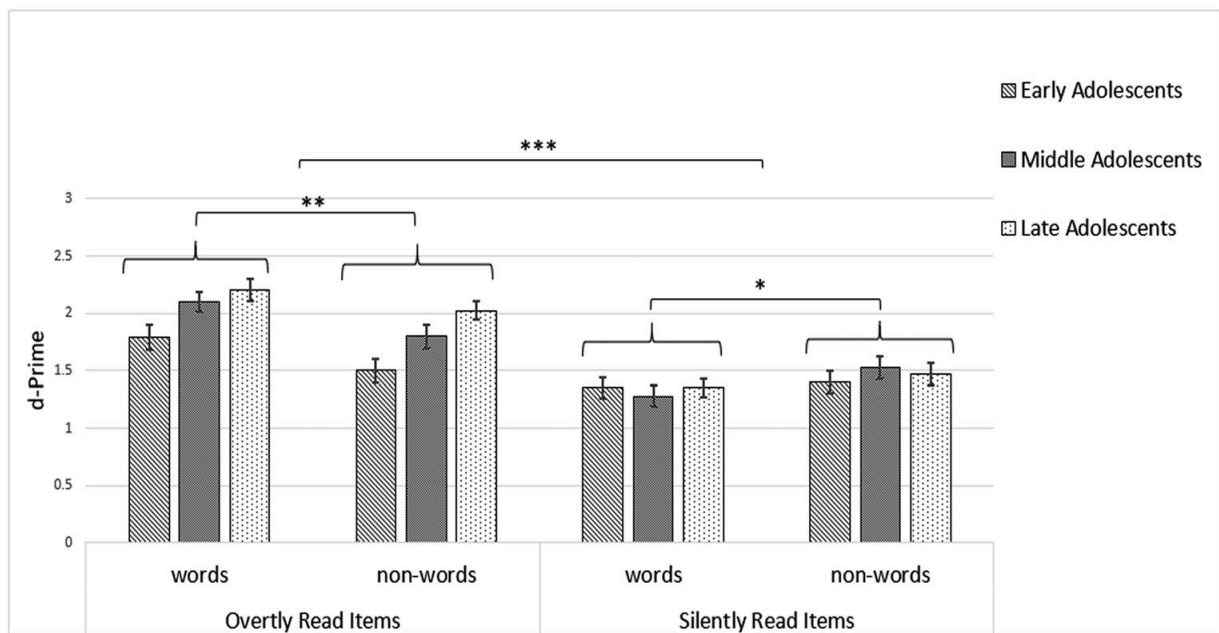
*p < 0.05
**p < 0.01
*** p < 0.001

Figure. 3.1. Means and standard errors for d-prime recognition scores in the early, middle and late adolescent groups

3.3.2.2. Self-monitoring performance

Table 3.3. presents the age-group means and standard deviations for the self-monitoring scores.

	Early Adolescents (n = 36, age = 13.04)		**Middle Adolescents** (n = 39, age = 15.13)		**Late adolescents** (n = 46, age = 17.02)	
Self-monitoring Test	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Self-monitoring errors Total (%)	32.83	10.22	31.01	11.79	33.22	11.92
Internalizing errors Total (%)	37.06	17.15	34.26	13.50	33.61	13.06
Externalizing errors Total (%)	28.59	13.63	27.81	15.54	32.82	16.79
Internalizing errors for words (%)	32.90	21.82	27.22	18.34	30.81	18.00
Internalizing errors for non-words (%)	41.22	25.86	41.19	16.77	36.42	21.19
Externalizing errors for words (%)	31.05	19.94	28.75	23.78	36.11	23.78
Externalizing errors for non-words (%)	26.13	18.80	26.87	21.74	29.52	20.31

Table 3.3. Age-group means and standard deviations for self-monitoring error percentages

Results of the mixed ANOVA conducted on the monitoring error scores for each type of item (words vs. non-words), with age-group as the between factor, revealed a main effect of monitoring error type (F (1, 118) = 8.53, p = 0.004), indicating a greater rate of internalizing than externalizing errors, independently of item type and age-group. Although no main effect for item type was found, the results demonstrate a significant interaction effect between monitoring error and item type (F (1, 118) = 10.36, p = 0.002), suggesting that the rate of monitoring errors affects differently each type of items. To follow-up this interaction effect, simple effects were analysed, showing a greater rate of internalizing errors for non-words than words (F (1, 118) = 12.58, p = 0.001) (Figure 3.2.).
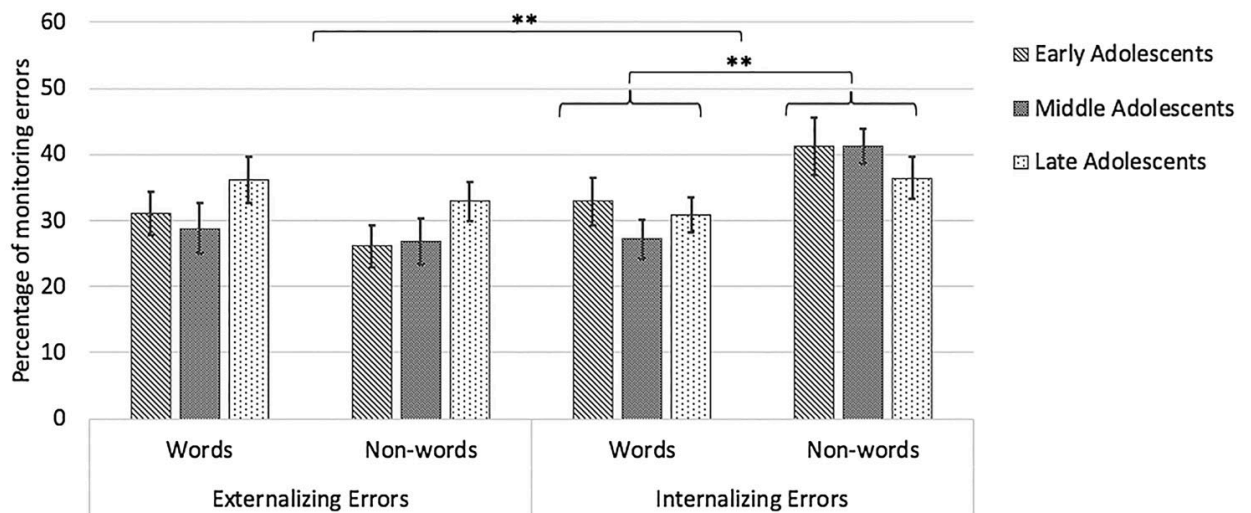
**p <0.01

Figure. 3.2. Means and standard errors for self-monitoring error scores in the early, middle and late adolescent groups

3.3.3. Associations between self-monitoring errors, schizotypal trait expression and impulsivity

Table 3.4. presents the correlations between all studied variables. Contrary to our hypothesis the cognitive-perceptual dimension of schizotypy was not associated with the rate of externalizing monitoring errors. Following Bonferroni correction, a significant positive correlation was found between the interpersonal SPQ dimension and the total percentage of self-monitoring errors (rs (121) = 0.22, p < 0.05). Neither the cognitive-perceptual nor the disorganization dimensions of schizotypy were associated with the total rate of self-monitoring errors in our sample.

Regarding impulsivity, none of the UPPS dimensions were associated with externalizing monitoring errors. Although the urgency UPPS dimension showed a negative correlation with internalizing monitoring errors (rs (121) = -0.22, p < 0.05), this relationship was not significant after Bonferroni correction was applied.

	1	2	3	4	5	6	7	8	9	10
1. Self-Monitoring Errors Total (%)	-	0.78**	0.74**	-0.01	0.22*	0.03	-0.11	-0.09	-0.09	0.01
2. Externalizing Monitoring Errors (%)		-	0.14	0.09	0.17	0.07	-0.04	-0.13	0.05	0.07
3. Internalizing Monitoring Errors (%)			-	-0.01	0.11	-0.06	-0.12	-0.02	-0.22*	-0.12
4. SPQ Cognitive-Perceptual				-	0.56**	0.70**	-0.02	0.07	0.34**	0.20*
5. SPQ Interpersonal					-	0.53**	-0.16	-0.04	0.33**	0.06
6. SPQ Disorganization						-	0.13	0.04	0.34**	0.22*
7. UPPS Lack of Premeditation							-	0.38**	0.02	0.08
8. UPPS Lack of Perseverance								-	-0.03	0.07
9. UPPS Urgency									-	0.24**
10. UPPS Sensation Seeking										-

Note: *$p < 0.05$, **$p < 0.01$

Table 3.4. Correlations between self-monitoring error percentages, SPQ and UPPS dimensions (before Bonferroni correction applied)

To confirm the correlation results presented in table 3.4, Bayesian statistics (BF_{10} and BF_{01}) of the correlations were also obtained. First, in line with the findings of the Spearman's correlations, Bayesian analyses showed evidence in favour of the association between the interpersonal dimension of the SPQ and the total percentage of self-monitoring errors ($BF_{10} = 3.32$), suggesting that the data were over 3 times more likely under the 2-sided alternative than the null hypothesis. In contrast, Bayesian analyses showed stronger evidence in favour of the null hypothesis when examining the correlation of the total rate of self-monitoring errors with the cognitive perceptual ($BF_{01} = 8.34$) and disorganization ($BF_{01} = 8.08$) dimensions of the SPQ. Furthermore, Bayesian analyses found more evidence for the null hypothesis when considering the correlation between the cognitive-perceptual dimension of the SPQ and the rate of externalizing self-monitoring errors for speech ($BF_{01} = 5.51$).

In terms of the associations between self-monitoring errors and impulsivity, results of Bayesian analyses showed stronger evidence in terms of the null hypothesis when examining the correlations of externalizing errors with the lack of premeditation ($BF_{01} = 7.45$); lack of perseverance ($BF_{01} = 2.96$); urgency ($BF_{01} = 7.28$) and sensation seeking ($BF_{01} = 4.94$) UPPS scales. Similarly, Bayesian analysis showed stronger evidence in terms of the null hypothesis when examining the correlations of internalizing self-monitoring errors with the lack of premeditation ($BF_{01} = 7.45$); lack of perseverance ($BF_{01} = 2.96$); urgency ($BF_{01} = 7.28$) and sensation seeking ($BF_{01} = 4.94$) UPPS scales. Finally, results of the Bayesian correlation confirmed the main analyses (prior to Bonferroni correction) showing a significant negative association between internalizing self-monitoring errors and the urgency dimension of the UPPS ($BF_{10} = 3.36$).

3.4. Discussion

The current study assessed the effects of age, cognitive effort and personality dimensions pertaining to schizotypy and impulsivity on self-monitoring for speech in community adolescents. The study employed a task that yields a recognition score and two self-monitoring error scores (internalizing and externalizing) for two types of items involving different levels of cognitive effort (words: low cognitive effort; non-words: high cognitive effort).

3.4.1.   Effects of age and cognitive effort on recognition performance

First, an age-group difference was found for recognition, suggesting that the early-adolescent group displayed worse recognition scores compared to middle and late adolescents. This is in line with developmental studies in adolescence, suggesting that memory for recognition improves with age (Murre et al., 2013). Closer inspection showed that the specific age-effect

was particularly pronounced for overtly-read items. This may suggest that younger adolescents do not utilize as effectively as their older peers the sensory-motor information that accompanies overt reading (i.e. articulation, sounds), which during retrieval can help discriminate between previously presented and new items.

Furthermore in accordance with the study hypothesis, results showed that silently-read items were less accurately recognized compared to overtly-read items. According to Johnson et al. (1993) the lack of sensory-motor signals during memory encoding for silently-read items may deem them less distinctive than overtly-read items, thus less accurately recognized. Interestingly, the type of reading affected differently the recognition scores for each type of item. Under overt reading, word items that require low cognitive effort were more accurately recognized than high cognitive effort non-word items. Conversely, under silent reading, participants displayed better recognition for non-word than word items. These findings are in line with the source-monitoring framework, which states that both recognition and source memory are influenced by changes in sensory-motor precision and cognitive activity during encoding (Johnson, 2006). The finding that under overt reading, adolescents recognize better word than non-word items may indicate that the encoding of sensory-motor signals yields a recognition advantage for low cognitive effort material. Accordingly, better recognition for non-word than word items under silent reading suggests that the absence of sensory-motor cues during encoding generates a memory advantage for high cognitive effort material.

3.4.2. Effects of age and cognitive effort on self-monitoring performance

Contrary to the current hypothesis and also in contrast to recognition performance, the present analyses failed to show an age-group effect for self-monitoring in adolescence. Indeed, research suggests that source memory is separate from recognition memory, and the two follow different developmental trajectories (Foley et al., 1983; Sussman, 2001). One possibility is that the capacity to distinguish between silent and overt speech is established prior to adolescence,

during childhood development (Sussman, 2001). Another possibility is that the task used is not sensitive enough to capture the developmental maturation that occurs in self-monitoring during adolescence (Lagioia et al., 2011).

Overall, results indicate that non-clinical youths are more prone to confuse overt speech as silent, while cognitive effort contributes to the overall rate of self-monitoring misattributions. Importantly, the current sample presented a specific pattern of self-monitoring errors for high cognitive effort items. In line with the hypothesis, non-word items led to more internalizing misattributions (confusing overtly-read items as silently-read). This is consistent with the source-monitoring framework, which states that conditions that increase the cognitive effort for the encoding of overt speech also increase its similarity to silent speech, thus making it difficult to discriminate between the two (Johnson et al., 1993). Another possibility is that increased cognitive effort reduces the encoding of sensory-motor cues that typically accompany overt speech, thus engendering uncertainty about the "realness" of overtly-read items (Morosan et al., 2018). While schizophrenia sufferers (Bentall et al., 1991) and high risk subjects (Debbané et al., 2010; Larøi et al., 2004) display externalizing source-monitoring misattributions under high cognitive effort, typically developing adolescents demonstrate the opposite pattern.

3.4.3. Associations of self-monitoring errors with personality traits of schizotypy and impulsivity

The present analyses suggest that self-monitoring errors are not associated with cognitive-perceptual or disorganization manifestations of schizotypy in adolescence. However, a significant positive association was found between the interpersonal dimension of schizotypy and the total rate of self-monitoring errors.

The lack of association between positive schizotypy and self-monitoring was unexpected. Most studies that have examined the relationship between schizotypy and self-

monitoring, albeit in adult samples, have reported associations with positive manifestations (Humpston et al., 2017; Peters et al., 2007). However, contrary to the current study that focused on speech-monitoring, self-monitoring studies in schizotypy have primarily used action-monitoring tasks (Humpston et al., 2017; Peters et al., 2007). Given that kinaesthetic actions entail richer sensory-motor and spatiotemporal traces than those involved in the production of overt speech, action-monitoring may constitute a relatively easier process than speech-monitoring (Collignon et al., 2005). Therefore, action-monitoring paradigms may be more sensitive than speech-monitoring tasks in capturing associations with cognitive-perceptual manifestations. Indeed, positive schizotypy has been consistently linked with difficulties in the monitoring of sensory-motor signals for both self-generated and other-produced physical actions (Itaguchi et al., 2018; Lemaitre et al., 2016). Furthermore, the only study that assessed both action- and speech-monitoring in a sample of healthy adults, reported that positive schizotypy was associated only with the former (Humpston et al., 2017).

The association between the interpersonal dimension of schizotypy and the total rate of self-monitoring misattributions in adolescence is a novel finding. Indeed, most self-monitoring studies in non-clinical samples have focused on cognitive-perceptual manifestations, either by using measures that assess hallucination and delusion-like phenomena, or by undertaking group comparisons according to the presence of these symptoms (Peters et al., 2007; Collignon et al., 2005). Contrary to these, Sahakyan and Kwapil (2016) assessed source-monitoring for verbally-presented material in a sample of healthy adults by comparing high scorers in positive schizotypy to those scoring high in negative schizotypy. They reported that the negative schizotypy subgroup exhibited significantly more source-monitoring errors compared to the positive schizotypy subgroup, indicating that negative rather than positive schizotypy involves impairments in source memory for verbally-presented material (Sahakyan & Kwapil, 2016). The current findings add to this literature, suggesting that the link between negative schizotypy

and source-monitoring errors may already be present during adolescence. One possibility is that impaired self-monitoring interferes with young people's ability to function within social situations, leading them to withdraw from interpersonal contact. Alternatively, because the cognitive processes that underpin the awareness of the self commonly develop within interpersonal relationships with others (Fonagy et al., 2002), reduced interpersonal contact in the context of negative schizotypy may contribute to undermine the elaboration of self-monitoring during adolescence (Debbané & Barrantes-Vidal, 2014). Given that negative schizotypy is one of the strongest distal predictors of clinical psychosis (Debbané et al., 2014), future investigations can prospectively examine whether exacerbations of self-monitoring misattributions in the context of negative schizotypy represent markers of psychosis vulnerability.

Regarding impulsivity, no associations were found between UPPS facets and self-monitoring. It has previously been argued that impairments in self-monitoring may contribute to limit incarcerated adolescent's insight about the negative consequences of their behaviours (Morosan et al. 2018). The lack of associations between self-monitoring and impulsivity in the current study suggests that this is not so for typically developing youths. Importantly however, self-report measures of impulsivity do not always correlate with behavioural assessments of impulse control, measured via response-inhibition tasks (Wilbertz et al., 2014). Thus, despite the lack of association with trait impulsivity, future studies can examine whether self-monitoring errors in adolescence are linked to momentary difficulties in inhibitory control.

3.4.4. Limitations and conclusions

The results of the current study should be interpreted in light of certain limitations. First, the data were derived from a relatively small sample and no prior power analyses were conducted to determine the sample size needed to detect significant effects. Thus, it remains possible that further associations could have emerged with a larger sample. It must be noted however that to

further explore whether the results of the main analyses represented true effects as opposed to non-conclusive findings and to reduce the probability of type II errors, Bayesian statistics were obtained and these provided support in favour of the study's findings. Nonetheless, future studies should test the results of the current study with larger samples. Second, the analyses performed were cross-sectional, and longitudinal investigations are needed to establish causal links between negative schizotypy and self-monitoring. Third, it must be noted that the mixed ANOVAs performed to statistically analyse the effects of the two within-subject factors on recognition (i.e. reading type and item type) and self-monitoring performance (i.e. monitoring error type and item type) could not account for individual differences between the participants included in the sample. The current study could have benefited by utilizing multilevel modelling statistical approaches to incorporate subject-related random effects to the within-subjects effects, thus account for individual differences in participant's recognition and self-monitoring performance under the different experimental conditions. Furthermore, it must be noted that in the current study the different SPQ scales correlated quite highly with each other, as well as with the urgency scale of the UPPS. Thus, it is possible that rather than solely capturing schizotypal traits, the SPQ may have, at least in part, also tapped into more global personality traits characterised by distress and affective dysregulation, such as neuroticism (Grant et al, 2018; Gross et al., 2014). This could also explain the observed correlation between SPQ scales and the urgency dimension of the UPPS, a factor characterised by high neuroticism (Gross et al., 2014)

Next, a number of methodological limitations pertaining to the specific experimental design used to test self-monitoring also need to be highlighted. First, although the decision to include non-words to create a high cognitive effort condition was based on previous evidence suggesting that their reading entails more effortful processing and stronger neural activation than that of words (Price & Devlin, 2011; Taylor et al., 2013), it cannot be assumed with

certainty that this effect was produced in the current sample. Indeed, the current study did not involve an objective measurement of cognitive effort. As such, it remains possible that the observed differences in recognition and self-monitoring performance between word and non-word items may have reflected the effects of factors other than cognitive effort (e.g. difficulty in reading and memorizing non-words compared to words). Relatedly, the different characteristics involved in the reading of words and non-words may have produced other confounding effects that could have impacted the results of the study. For instance, given that the reading of non-words is more difficult than that of familiar words, participants may have assumed that the experimenter expected them to show better recognition and self-monitoring performance for the latter compared to former, thus adjusting their effort accordingly. Although the first part of the task was introduced as a reading and pronunciation exercise and participants were not aware that a memory test would follow, we cannot discount the possibility that some participants may have adjusted their performance during the recognition and self-monitoring tests on the basis of their perceived demand characteristics of the experiment. Furthermore, although participants were asked to read all test items presented in the task, it cannot be assumed with certainty they did so for items they had to read silently. Indeed, contrary to overtly-read items, the reading of silent items was not observed by the experimenter, thus it remains possible that some participants may not have read the items silently, particularly as it pertains to the more difficult non-word items. Finally, the current study did not account for differences in the emotional valence and semantic meaning of the word items used. This is important as previous research indicates that the emotional valence and semantic processing of word items may play a part in the association between externalizing self-monitoring misattributions and hallucination-like experiences (Larøi et al., 2004; Sugimori et al., 2011).

Despite these limitations, this study is the first to examine self-monitoring for speech in typically-developing adolescents and the nature of its associations with personality traits

pertaining to schizotypy and impulsivity. The present analyses highlight the importance of cognitive effort and negative schizotypy on the level of self-monitoring confusions across adolescence. From an early prevention standpoint, findings suggest that difficulties in the monitoring of self-generated mental events may represent psychotherapeutic targets to attenuate trait-risk for schizophrenia in youths presenting with negative schizotypal features.

Chapter 4

Reality-monitoring for speech and its links to age, emotional valence and schizotypal trait expression during adolescence

4.1. Introduction

Contemporary clinical research in psychosis has sought to identify the cognitive processes that underpin the development and maintenance of the clinical manifestations characterizing the illness. Because of their high prevalence, as well as their deleterious effects on functional outcomes, positive symptoms of psychosis, such as hallucinations and delusions have been the primary focus of these investigations. As mentioned in section 1.6.3, a number of conceptual accounts postulate that the cognitive-perceptual aberrations observed in psychotic disorders may be underpinned by impairments in *reality-monitoring* – that is the cognitive capacity to discriminate between information generated by oneself and information generated by another agent. For instance, hallucinations have been conceptualised as stemming from misattribution of one's self-generated internal content (e.g. inner voice) to a source external to the self (Allen et al., 2007), while delusions of control have been attributed to difficulties in monitoring the initiation of self-generated motor actions, whose control is instead attributed to external agents (Blakemore et al., 2003).

Reality-monitoring research in people suffering with psychosis has consistently shown that confusions in the identification of self-generated material play a critical role in the development and maintenance of positive psychotic manifestations. For instance, patients suffering with schizophrenia commonly fail to recognise their own pre-recorded speech when slight pitch modifications are experimentally introduced in the recordings (Johns et al. 2001). Similarly, research has shown that patients suffering with the illness experience difficulties in

discriminating in memory between speech generated by themselves and speech generated by another agent (Bental et al., 1991; Brébion et al., 2000; Vinogradov et al., 2008; Morrison and Haddock et al., 1997). More specifically, the majority of studies suggest that patients suffering from the illness display a tendency towards misattributing the source of self-generated speech to external agents (Bentall et al., 1991; Brébion et al., 2000; Johns et al., 2010; Johns et al., 2006). Importantly, this pattern of externalizing reality-monitoring misattribution appears to be particularly pronounced among patients reporting hallucinations (Bentall et al., 1991; Brébion et al., 2000), while some evidence also suggest links with delusional ideation (Johns et al., 2001; Franck et al., 2001).

Another line of research indicates that reality-monitoring impairments are also observable, at an attenuated level, in the prodromal and premorbid stages of psychosis, as well as among non-affected first-degree relatives of patients suffering from the illness, suggesting that they may play a critical role in the developmental unfolding of the clinical forms of the illness. For example, Johns et al (2010) showed that similarly to patients with psychosis and in contrast to non-clinical controls, CHR adults misattribute audio-recordings of their own speech to external sources when the pitch of the sound is experimentally manipulated. Interestingly however, no linear associations were reported between the rate of reality-monitoring misattributions and specific prodromal symptoms. Action simulation studies among first-degree relatives of patients with psychosis also suggest that non-clinical individuals at genetic-risk for psychosis experience confusions in discriminating between self-generated and experimenter-produced motor actions (Versmissen et al., 2007; Hommes et al., 2012). The genetic-risk groups in these studies show lower error rates compared to patients with psychosis, but significantly greater error rates compared to non-clinical controls, suggesting a dose-response of psychosis-risk on reality-monitoring confusions. Furthermore, within groups of first-degree relatives of patients with psychosis a significant positive association has been

found between the total rate of reality-monitoring misattributions and delusional ideation (Versmissen et al., 2007), as well as scores on measures of positive schizotypy (Hommes et al, 2012).

In non-clinical samples from the community, studies suggest that individuals who score high (e.g. scoring within the 25th upper percentile) on measures of positive schizotypy, hallucination-proneness and delusional ideation display significantly more externalizing reality-monitoring errors in the attribution of self-generated motor actions, as well as in the recollection of self-generated speech, compared to groups with low scores on these measures (Versmissen et al., 2007; Larøi et al., 2004; Allen et al., 2007). Although suggestive of a relationship between reality-monitoring and sub-clinical signs of psychosis, these studies involve adult participants either within or beyond the critical period of illness onset. Furthermore, the majority of reality-monitoring studies in adults focus on cognitive-perceptual manifestations, either by using measures that assess hallucination and delusion-like phenomena, or by undertaking group comparisons according to the presence of these symptoms. From an early prevention standpoint, important insights can be gained by examining reality-monitoring and its associations with schizotypal personality dimensions during earlier developmental stages.

Adolescence represents a critical developmental period characterised by a number of biopsychosocial changes, which generate increased emotional, interpersonal and cognitive stress load that confers increased vulnerability to a number of mental health problems (Steinberg and Morris, 2001). Indeed, the first signs of psychopathological expression often manifest during adolescent development and adolescence represents a key period during which schizotypal manifestations become clinically relevant (Gooding et al., 2005). To date, only a limited number of studies have analysed the associations between schizotypal trait dimensions and source-monitoring difficulties in adolescent samples. In an action-monitoring study

(Debbané et al., 2009) a sample of non-clinical adolescents was presented with a series of action statements (e.g. "open the door") and were asked to either (1) visualize themselves performing the actions; (2) visualize the experimenter performing the action; or (3) mentally repeat the action statement without visualization of the performer. When asked to recall in memory the source of the action-statements results showed that positive schizotypy in the non-clinical youth sample was significantly correlated with increased rates of source confusions in discriminating between imagine-experimenter and mentally-repeat actions (i.e. discriminating between the external conditions of the task that did not involve participants visualizing themselves performing the action). In a speech-monitoring study, Debbané et al (2010) found that in a group of adolescents with 22q11.2 deletion syndrome, a neurogenetic condition conferring high-risk for schizophrenia, positive schizotypy was significantly associated with the tendency to misattribute in memory word items they had read silently as overtly read (i.e. externalizing self-monitoring misattribution).

Contrary to these findings, results of the second study presented in the current thesis (also published in Salaminios et al., 2020b), which applied the same speech-monitoring task in a sample of non-clinical youths did not find an association between the positive dimension of schizotypy and self-monitoring errors. Instead, a significant positive correlation was found between scores on the negative dimension of schizotypy and the total rate of self-monitoring errors (i.e. sum of externalizing and internalizing errors). Thus, in contrast to studies in clinical psychosis and non-clinical adult samples, these findings suggest that confusions between imagined and real events in adolescence may be linked to interpersonal aspects of schizotypy (Salaminios et al., 2020b). One possibility is that the associations commonly observed between source-monitoring difficulties and cognitive-perceptual aberrations in clinical and hallucination-prone adult samples, may be preceded by more subtle associations with interpersonal features of schizotypy during adolescence. Importantly though, no study to date

has directly examined the associations between schizotypal trait dimensions and the capacity to discriminate in memory between self-generated and other-produced speech in non-clinical adolescents. Research on adolescent schizotypy can further our understanding regarding the early course of reality-monitoring impairments during the premorbid stages of schizophrenia-spectrum illnesses, prior to the development of clinically relevant symptoms.

As discussed above, a number of studies have shown presence of reality-monitoring confusions in samples of patients suffering with psychosis and among non-clinical individuals in the context of schizotypal trait expression. However, the underlying factors that underpin these confusions remain incompletely understood. The emotional valence of stimuli has been investigated as one contributing factor that may disrupt reality-monitoring performance. According to the source-monitoring framework (Johnson et al., 1993), judgments about the source of memory information are primarily based on evaluations pertaining to the characteristics of the memory events themselves, such as the contextual (spatial/temporal) and perceptual characteristics (visual/auditory) available during memory encoding and retrieval. While some studies also suggest that the amount and clarity of emotional information stored in memory may also be used as cues to help identify the source of the event (Johnson and Suengas, 1989), other studies indicate that increased emotional involvement during memory encoding can disrupt the processing of perceptual and contextual information that are critical in sustaining accurate source-monitoring discriminations. For instance, Suengas and Johnson (1988) asked participants to engage in a number of activities (e.g. writing a letter) and subsequently think about either factual (how things looked), or affective aspects (how they felt) related to these activities. When asked to recall various phenomenal characteristics pertaining to these events, results showed that thinking about affective aspects of the events during encoding reduced the availability of memories for perceptual information. On the basis of these findings and in order to directly test whether emotional involvement affects source-memory,

Johnson et al (1996) presented participants with audio-recordings of two people (one male and one female) reading various statements under two emotional-focus encoding conditions. Specifically, while attending to the recordings of the two speakers, participants were asked to either focus on how they felt about what was being said (self-focus condition) or focus on how they thought the speakers felt (other-focus condition). In the test phase participants were presented with a list containing the statements the two speakers made in the first part of the task and were asked to recall the source of these statements (i.e. female or male speaker). Results showed that focusing on one's own feelings about what was being said led to reduced source-monitoring accuracy relative to focusing on the speakers' feelings. Furthermore, within the self-focus condition, the rate of source-monitoring errors was higher for the statements with the highest emotional load. According to Johnson et al (1996) focusing on how one feels about an event may activate a number of self-reflective processes (i.e. considerations of one's beliefs, opinions, previous experiences), which may in turn reduce the externally-oriented processing of the perceptual, contextual and semantic information that commonly support accurate source-monitoring discriminations.

In accordance to the source-monitoring framework, most studies to date suggest that emotional valence of events appears to exacerbate reality-monitoring confusions in clinical psychosis samples and hallucination-prone non-clinical adults. For example, Morrison and Haddock (1997) found that patients suffering with auditory hallucinations made more reality monitoring errors for speech (misattributing self-generated words to the experimenter and vice versa) for emotionally charged material (positive and negative word items) compared to healthy controls. Further analyses showed that patients experiencing hallucinations made more externalizing misattribution errors for positively-charged words, while more internalizing misattributions were found for negatively-charged words. Similarly, Larøi et al (2004) found that a group of non-clinical adults prone to experience hallucinations tended to misattribute to

the experimenter word-items they had generated themselves when the words were emotionally charged. Post-hoc analyses showed that hallucination-prone adults committed more reality-monitoring misattributions for negatively charged than positively charged words. These findings suggest that the emotional valence of material may exacerbate reality-monitoring errors among both clinical and non-clinical individuals experiencing hallucinatory-phenomena. Given the evidence for increased emotional arousal among patients with schizophrenia and hallucination-prone individuals (Larøi et al., 2004), one possibility is that emotionally salient material may evoke a significant degree of arousal in individuals who experience hallucinations, thus further perturbing their reality-monitoring abilities.

Despite this evidence, no study to date have examined the effects of emotional valence on reality-monitoring during adolescence. This is surprising as adolescence represents a critical developmental period, characterised by a number of biopsychosocial changes and interpersonal challenges (e.g. hormonal changes, initiation of romantic relationships, reduced reliance on parental support, new academic pressures) that significantly impact on the capacity to regulate emotions (Garnefski et al., 2001). Given that the first signs of clinical psychosis often emerge between adolescence and young adulthood, from an early prevention approach studying the effects of emotional valence on reality-monitoring among young people is clinically relevant.

Interestingly, little is known at present about the developmental elaboration of reality-monitoring processes during adolescence. In contrast to self-monitoring, which appears to improve between childhood and adulthood, developmental research on reality-monitoring suggests that the capacity to differentiate in memory between self-generated and other-produced actions is established in early childhood. For instance, Folley et al (1983) found that 6-year-old children performed as good as 17-year-olds in discriminating between memories of what they had said earlier from memories of what they had said another person saying. Similarly, Folley and Johnson (1985) found that 6- and 9-year olds performed as well as adults

in distinguishing between enacted and imagined physical actions. Therefore, despite neuroimaging data showing shared neural activation patterns during reality- and self-monitoring evaluations (Lagioia et al, 2011), evidence from developmental research point towards different developmental trajectories between the two. Given the lack of differences in reality-monitoring performance between young children and adult participants, no study to date has examined the potential effects of age on reality-monitoring during adolescence. It must be noted however that previous studies examining differences in reality-monitoring between children and adults have primarily used experimental-tasks that do not include emotionally charged material. This is important given that, as mentioned earlier, the emotional valence of material appears to affect reality-monitoring performance. Furthermore, the biopsychosocial changes that take place in adolescence commonly lead to increased emotional stress load, while at the same time promoting the elaboration of neurobiological and psychological processes that sustain the regulation of emotion activation (Garnefski et al., 2001; Tanzer et al., 2020). As such adolescence may represent a critical developmental period for the specialization of reality-monitoring evaluations pertaining to emotionally charged material.

The present study will examine reality-monitoring in community adolescents and young adults using a task that assesses the capacity to discriminate in memory between self-generated and experimenter-produced speech under different levels of emotional valence (Larøi et al., 2004). The main aims of the study are to assess the effects of age and emotional valence on reality-monitoring for speech in adolescence, as well as its associations with schizotypy dimensions.

Given that previous research indicates that young children exhibit comparable performance in reality-monitoring with adults (Folley et al., 1983; Folley and Johnson, 1985), it is hypothesised that adolescents will exhibit similar rates of reality-monitoring errors to young adults. However, in line with developmental research on emotion processing (Ahmed et

al., 2015), it is also hypothesised that adolescents will exhibit a higher rate of reality-monitoring errors for emotionally charged material compared to young adults. Furthermore, in line with the source-monitoring framework, it is expected that the current sample will demonstrate increased rates of reality-monitoring errors for emotionally charged compared to emotionally neutral material, as the processing of the former commonly disrupts the encoding of contextual and perceptual information (Suengas and Johnson, 1988; Johnson et al., 1996).

In terms of the associations with schizotypal traits, following previous reality-monitoring research in patients with psychosis and non-clinical adults (Bental et al., 1991; Brébion et al., 2000; Morrison and Haddock, 1997; Larøi et al., 2004), it is hypothesised that the rate of externalizing misattributions for speech will be positively associated with the cognitive-perceptual dimension of schizotypy. Because the majority of reality-monitoring studies have only investigated cognitive-perceptual manifestations and no study to date has examined reality-monitoring for speech and it's links to schizotypal trait expression in community adolescents, the current study also seeks to explore whether higher schizotypy scores in the interpersonal and disorganization dimensions, will be associated with increases in the total rate of reality-monitoring misattributions (used in the current study as a measure of general self-monitoring performance). In accordance with results from study 2 (Chapter 3) of the current thesis, which found links between self-monitoring and negative schizotypy in adolescence (Salaminios et al., 2020b), it is hypothesised that the interpersonal dimension of schizotypy will also be associated with the total rate of reality-monitoring errors in the current sample.

4.2. Methods

4.2.1. Participants and procedure

A convenience sample of 84 community adolescents and young adults was recruited through written advertisements in public schools in the city of Geneva, Switzerland. No prior power analysis was performed to determine the size of the sample Inclusion criteria were age (12-20 years) and fluency in French. All participants were screened for cognitive impairment. Those with scores below a standard score of 7 in the Block Design subtest of the Wechsler Intelligence Scale (Wechsler, 1955) were excluded from the analyses (n = 8). The final sample consisted of 76 adolescents and young adults (35 females, 41 males; M_{age}= 17.00). None of the participants suffered from past/present psychiatric, neurological or neurogenetic disorders. Furthermore, none of the participants of the current study participated in studies 1 or 2 of the current thesis. Written informed consent was obtained from all participants and legal guardians of those under the age of 18. Ethical approval for the study was granted by the Swiss Ethics Committee on research involving humans (number: 2018-00251).

4.2.2. Measures

4.2.2.1. Self-report measures

The Schizotypal Personality Questionnaire (SPQ-French version, Dumas et al., 2000) measures schizotypal traits and their dimensions, yielding three factor scores: cognitive–perceptual (unusual perceptual experiences, suspiciousness, ideas of reference); interpersonal (social anxiety, constricted affect, lack of close friends); and disorganization (odd speech, odd behaviour). The French version of the SPQ has shown good reliability (Cronbach's alpha = 0.91) (Dumas et al., 2000) and has been validated for use with francophone adolescents (Badoud et al., 2011). The SPQ in the present study also showed good reliability (Cronbach's alpha = 0.89).

4.2.2.2. Reality-monitoring task

The reality-monitoring task consisted of the experimenter verbally presenting participants with a list of 30 words, 10 of which were positive, 10 negative and 10 neutral (see appendix 5). The word-items used for this task were taken from a previous study (Larøi et al., 2004) for which words were rated by healthy subjects according to emotional valence (i.e. positive, neutral, negative), length, and their degree of agreeability/ disagreeability (see below information about the selection of word items).

Selection of word items

Larøi et al., (2004) followed four separate steps to select the word items for the reality-monitoring task. In the first step they chose a set of highly positive, negative and neutral words from a previous study (Brain et al, 2000) in which words were rated on a 1-7 Likert-type scale according to their degrees of dysphoria, anxiety and agreeability, resulting in groups of negative (high degrees of dysphoria and anxiety), positive (high degrees of agreeability) and neutral (low degrees of agreeability, anxiety and dysphoria). From these a set of words were chosen according to their length (number of syllables and letters) resulting in 229 total items (83 positive, 72 negative and 74 neutral). In the second step Larøi et al., (2004) presented these words in random order to a another group of healthy subjects (n = 20) who rated each word on its degree of agreeability/disagreeability on a 1 (high disagreeability) to 7 (high agreeability) Likert scale. The then selected the items with scores between 6 and 7 (positive), between 1 and 2 (negative) and between 3 and 5 (neutral), resulting in 124 words (43 positive, 39 neutral, 43 negative). In the next step, only the words that were between 650 and 900 milliseconds in length when spoken were kept, resulting in 90 words (30 positive, 30 negative and 30 neutral). In the final step, each word was orally presented to another group of healthy subjects (n =15) who were asked to respond with the first word that came to their mind. Reaction times for each

word were recorded. For each valence type (positive, negative, neutral), the five original words with the longest reaction time and the 5 words with shortest reaction time were kept, resulting in the final 30 words (10 for each valence type). Then 30 more words were selected to act as the distractor items, which were again controlled for emotional valence and length.

For each word presented by the experimenter, participants were asked to say the first word that came into their mind. Subjects were free to select any word they wanted as long as it did not consist of proper nouns or did not come from the same family of words as the item presented by the experimenter (i.e. work - worker). Participants were also asked not to repeat words that the experimenter or themselves had already said. After a 15-minute interval delay, a recognition sheet was handed out, containing 90 items (the 30 original items produced by the experimenter, the 30 items generated by the participant, and 30 new distractor items). In the test phase, participants had to indicate if each item from the list appeared in the first part of the task (recognition test), and attribute whether each recognized item was originally generated by the experimenter (other-generated) or by themselves (self-generated). The task yields two types of self-monitoring errors: externalizing errors (attributing word items generated by oneself to the experimenter), and internalizing errors (attributing word items generated by the experimenter to oneself); for three item types involving different emotional valence: positive, negative and neutral.

For the recognition test, signal detection theory (Stanislaw & Todorov, 1999) was used to estimate d' scores by subtracting the z score corresponding to total number of false alarms from the z score corresponding to the total number of hits.

The percentage of externalizing errors was calculated by dividing the total score for items generated by the participant but identified as generated by the experimenter, out of the total number of correctly recognized self-generated items. The percentage of internalizing errors was calculated by dividing the total score of items generated by the experimenter but

identified as generated by the participant, out of the total number of correctly recognized experimenter-generated items. Percentages of externalizing and internalizing errors were calculated overall, as well as for each item valence type. Finally, the total percentage of reality-monitoring errors was calculated as the sum of internalizing and externalizing errors for all items.

4.2.3. Statistical analyses

A one way ANOVA was first conducted on *d'* recognition scores with age-group as a between subject factor [13-16.20 years-of-age (early adolescents, n = 21), 16.20-18 years-of-age (late adolescents, n = 30), and 18-20 years-of-age (young adults, n = 25)].

A 2×3×3 mixed ANOVA was conducted on reality-monitoring error scores with two within subject factors: (reality-monitoring error type: externalizing vs. internalizing; and item valence type: negative vs. positive vs. neutral), and age-group as the between factor [13-16.20 years-of-age (early adolescents, n = 21), 16.20-18 years-of-age (late adolescents, n = 30), and 18-20 years-of-age (young adults, n = 25)]. Because male and female participants did not differ in their reality-monitoring scores, gender was not entered as a covariate in these analyses.

Because all SPQ scale scores violated the assumption of normality (Kolmogorov-Smirnov test, $p < 0.05$), Spearman's correlation coefficients were calculated to estimate their relationships with reality-monitoring errors (externalizing, internalizing and total rate of reality monitoring errors). To account for multiple comparisons and reduce the probability of type I errors, Bonferroni adjusted significance levels were calculated.

Finally, given that no prior power analysis was performed to determine the sample size necessary to detect significant effects and in order to explore whether the results of the correlations assessing the relations between SPQ scales and reality-monitoring errors represented true effects as opposed to non-conclusive findings, JASP version 0.9.2 was used

to obtain simple Bayesian statistics of the correlation analyses run (Nuzzo, 2017). In contrast to standard hypothesis testing based on *P* values, Bayesian statistics offer a direct measure of the strength of the evidence both in favour and against the study hypotheses that can support the interpretation of results and reduce the probability of type II errors. Specifically, Bayesian correlations provide a Bayes factor (BF_{10}) that represents the likelihood of the data under the null hypothesis (i.e. no correlation) divided by the likelihood of the data under the alternative (i.e. correlation present), so that BF_{10} values greater than 1 signal more evidence in favour of the alternative hypothesis compared to the null. Furthermore, Bayesian correlations provide another Bayes factor (BF_{01}) which represents $1/BF_{10}$, so that BF_{01} values greater than 1 signal more evidence in favour of the null hypothesis compared to the alternative. Given that both BF_{10} and BF_{01} contain the same information, for ease of interpretation the results section will only report the Bayes factor that is greater than 1.

4.3 Results

4.3.1 Descriptive results

Table 4.1. presents the descriptive results for the variables included in our analyses.

	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
Recognition Test						
Correct recognition/Hits raw score	53.42	4.00	40.00	59.00	-1.04	1.07
Correct recognition/Hits (%)	89.04	6.70	67.00	98.00	-1.04	1.07
False alarms raw score	1.55	1.96	0.00	12.00	2.51	9.92
d-prime total	3.03	0.55	1.81	3.96	-0.27	-0.60
Reality-monitoring Test						
Reality-monitoring errors Total (%)	6.08	4.92	0.00	22.72	1.14	1.58
Internalizing errors Total (%)	5.49	5.92	0.00	27.59	1.79	4.05
Externalizing errors Total (%)	6.57	6.15	0.00	22.22	0.63	-0.75
Internalizing errors for positive words (%)	7.46	9.29	0.00	42.86	1.24	1.74
Internalizing errors for negative words (%)	8.32	12.20	0.00	55.56	1.88	3.72
Internalizing errors for neutral words (%)	1.48	4.29	0.00	20.00	3.01	8.78
Externalizing errors for positive words (%)	6.67	8.47	0.00	33.33	1.16	0.80
Externalizing errors for negative words (%)	10.06	11.66	0.00	10.06	1.01	0.36
Externalizing errors for neutral words (%)	7.32	5.61	0.00	22.22	1.60	1.80
Schizotypal Traits (SPQ)						
SPQ Cognitive-perceptual dimension	7.32	6.78	0.00	27.00	0.95	0.02
SPQ Interpersonal dimension	5.18	4.25	0.00	17.00	0.94	0.10
SPQ Disorganization dimension	5.96	4.07	0.00	16.00	0.55	-0.28

Table 4.1. Means and standard deviations for each variable in the sample
Abbreviations: SPQ, Schizotypal Personality Questionnaire

Kurtosis and skewness of the SPQ scale distributions were all within the acceptable range (i.e. all values between -2.00 and 2.00) and no outliers were identified. Furthermore, the distribution of SPQ scores showed that the proportion of participants reporting elevated schizotypal traits (1.5 standard deviations above the mean) was comparable across the different SPQ scales (10.52% for the interpersonal [n = 8]; 10.52% for the cognitive-perceptual [n = 8] and 10.52% for the disorganization [n = 8] SPQ dimensions.

4.3.2   Reality-monitoring task results

4.3.2.1   Recognition performance

Table 4.2. presents the age-group means and standard deviation scores for correct recognition

(hits), false alarm and *d'* scores.

	Early Adolescents (n = 21 , age = 14.66)		Late Adolescents (n = 30 , age = 17.11)		Young Adults (n = 25 , age = 18.85)	
Recognition Test	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Correct recognition/Hits raw score	53.00	0.80	53.13	0.86	54.12	0.68
Correct recognition/Hits (%)	88.33	1.33	88.56	1.42	90.20	1.12
False alarms raw score	2.29	0.68	1.50	0.23	1.00	0.22
d-prime Total	2.87	0.13	2.99	0.09	3.20	0.11

Table 4.2. Age-group means and standard deviations for Hits, false alarm and d-prime scores

Results of the one-way ANOVA conducted on *d'* scores showed that there were no age-group

differences in recognition performance (F (2, 73) = 2.23, p = 0.12).

4.3.2.2   Reality-monitoring errors

Table 4.3. presents the age-group means and standard deviations for the reality-monitoring

error scores.

Self-monitoring Test	Early Adolescents (n = 21 , age = 14.66)		Late Adolescents (n = 30 , age = 17.11)		Young Adults (n = 25, age = 18.85)	
	Mean	SD	Mean	SD	Mean	SD
Reality-monitoring errors Total (%)	6.39	2.10	7.01	1.89	4.69	1.82
Internalizing errors Total (%)	5.53	1.29	5.49	1.23	5.46	1.02
Externalizing errors Total (%)	7.24	1.45	8.54	1.01	3.92	1.13
Internalizing errors for positive words (%)	5.44	1.81	6.83	1.56	9.92	2.13
Internalizing errors for negative words (%)	10.29	2.80	8.70	2.29	6.21	2.28
Internalizing errors for neutral words (%)	1.48	0.81	1.70	0.86	1.20	0.88
Externalizing errors for positive words (%)	10.02	2.44	6.00	1.52	4.67	1.22
Externalizing errors for negative words (%)	8.85	2.22	15.28	2.12	4.81	2.11
Externalizing errors for neutral words (%)	2.54	1.02	4.35	1.13	2.33	1.12

Table 4.3. Age-group means and standard deviations for reality-monitoring error percentages

Results of the mixed ANOVA conducted on the reality-monitoring error scores revealed a main effect of item emotional valence type (F (1, 73) = 28.14, $p < 0.001$), indicating a significantly greater rate of reality monitoring errors for both positive and negative word items compared to neutral items, independently of error type and age-group.

Furthermore, the results showed a significant interaction effect between item valence type and age group (F (4, 73) = 3.12, $p < 0.05$). To follow-up this interaction effect, simple effects were analysed, showing that the late adolescent group exhibited significantly more reality monitoring errors for negative items compared to the young adult group ($p < 0.05$), independently of error type.

Although no main effects were found for reality-monitoring error type, the results demonstrate a significant interaction effect between error type and age group (F (2, 73) = 3.20, $p < 0.05$). To follow-up this interaction effect, simple effects were analysed, showing that the late adolescent group exhibited significantly more externalizing reality monitoring errors compared to the young-adult group ($p < 0.01$), independently of item valence type.

Finally, results revealed a significant triple interaction effect between valence type, reality-monitoring error type and age group (F (4, 73) = 2.90, $p < 0.05$). Analyses of simple

effects showed that the early adolescent group exhibited more externalizing errors for positive items compared the young adult group ($p < 0.05$), while the late adolescent group exhibited more externalizing errors for negative items compared to both the early adolescent ($p < 0.05$) and young adult groups ($p < 0.01$) (Figure 4.1.).
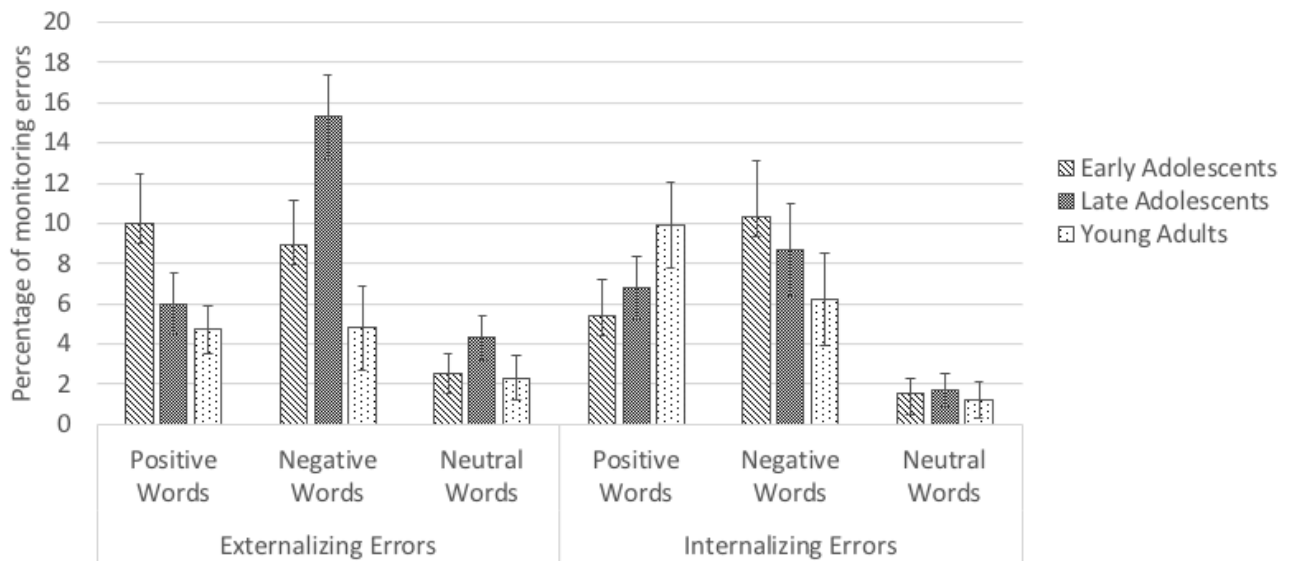
Figure. 4.1. Means and standard errors for reality-monitoring error scores (externalizing and internalizing) for three emotional valence conditions (positive, negative, neutral) in early adolescents, late adolescents and young adults

4.3.3 Associations between reality-monitoring errors and schizotypy dimensions

Table 4.4. presents the correlations between all studied variables. Contrary to the hypothesis the cognitive-perceptual scale of the SPQ was not associated with the rate of externalizing reality-monitoring errors. Neither the interpersonal, nor the disorganization dimensions of the SPQ were associated with externalizing or internalizing reality-monitoring errors. Similarly, no significant associations were found between the total rate of reality-monitoring errors and SPQ scales.

		1	2	3	4	5	6
1.	Reality-Monitoring Errors Total (%)	-	0.85[**]	0.74[**]	-0.04	-0.04	0.07
2.	Externalizing Monitoring Errors (%)		-	0.31[**]	-0.01	0.02	0.11
3.	Internalizing Monitoring Errors (%)			-	-0.13	-0.12	-0.05
4.	SPQ Cognitive-Perceptual				-	0.63[**]	0.75[**]
5.	SPQ Interpersonal					-	0.52[**]
6.	SPQ Disorganization						-

Note: $*p < 0.05$, $**p < 0.01$

Table 4.4 Correlations between reality-monitoring error percentages and SPQ dimensions

To confirm the results presented in table 4.4, Bayesian statistics (BF_{10} and BF_{01}) of the correlations were also obtained. In line with the findings of the Spearman's correlations, Bayesian analyses showed evidence in favour of the null hypothesis when examining the correlation between the rate of externalizing reality-monitoring errors and the cognitive perceptual scale of the SPQ ($BF_{01} = 6.60$). Furthermore, Bayesian analyses found more evidence for the null hypothesis compared to the alternative when considering the correlations of the total rate of reality-monitoring errors with the cognitive-perceptual ($BF_{01} = 6.12$), interpersonal ($BF_{01} = 6.30$) and disorganization ($BF_{01} = 4.95$) SPQ scale scores.

4.4 Discussion

The current study assessed the effects of emotional valence and age on reality-monitoring for speech, as well as its association with personality dimensions pertaining to schizotypy in community adolescents and young adults. The task employed yields two reality-monitoring error scores (internalizing and externalizing) for three types of word-items involving different levels of emotional valence (positive, negative, neutral).

4.4.1 Effects of emotional valence and age on reality-monitoring performance

First of all, in line with the hypotheses, results of the current study suggest that adolescents and young adults committed significantly more reality-monitoring errors for emotionally charged stimuli (positive and negative words), compared to neutral material. This finding is in consistent with previous research, suggesting that the activation of one's emotional arousal during memory encoding reduces reality-monitoring accuracy (Johnson et al., 1996). In accordance to the source-monitoring framework (Johnson et al), the negative effect of emotion on reality monitoring performance identified in the current study may be interpreted by suggesting that in adolescents and young adults, emotionally salient words evoke a significant degree of arousal, which in turn perturbs the encoding of contextual characteristics that normally sustain accurate source-monitoring attributions. Contrary to patients suffering with psychosis and non-clinical hallucination-prone adults, emotionally-charged material did not lead to more externalizing errors (misattributing self-generated items as experimenter-produced) in the current non-clinical sample. Rather, when the sample was taken as a whole, emotionally salient items led to a more general confusion in the differentiation between self and non-self-cues that involved both externalizing and internalizing errors.

In line to the hypotheses the present analyses failed to show a main effect of age for the total rate of reality-monitoring errors in adolescents and young adults. This is also in accordance with developmental research suggesting that the capacity to distinguish between self-generated and externally-produced speech is established prior to adolescence, during early childhood (Folley and Johnson, 1985; Folley et al., 1983). Importantly however, adolescents and young adults differed in the rate of errors they committed when the material were negatively charged. More specifically, results show that the late adolescent group made more reality-monitoring errors than young adults for negatively charged word-items. Indeed,

neuroimaging data suggest that although the prefrontal cerebral networks that sustain reality-monitoring are already in place prior to adolescence, their functional specialization and accuracy may be further elaborated during adolescent development (Lagioia et al., 2011). This is interesting given that the developmental elaboration of the prefrontal networks implicated in reality monitoring during adolescence coincides with ongoing functional maturation in the prefrontal cortex as it pertains to its activity in the regulation of negative emotions (Young et al., 2019; Ahmed et al., 2015). Thus, one possibility is that when adolescents undertake reality-monitoring judgments for material that evoke negative emotional arousal, the neural demands imposed on the prefrontal cortex may exceed its functional maturation, leading to disruptions in reality-monitoring performance. Importantly though, in contrast to late adolescents, no differences were found in reality monitoring-performance for negatively charged items between the early adolescent and young adult groups. One possibility is that the initiation of various emotional challenges specifically during late adolescence, such as the initiation and increasing importance of peer and romantic relationships, reduced dependence on family support, and new academic pressures (Casey et al., 2010), may lead to heightened emotional reactivity towards negative material, thus contributing to disruptions in reality-monitoring performance.

Furthermore, although no main effects of age or reality-monitoring error type (externalizing vs internalizing) were found, a significant interaction effect emerged between age-group and reality-monitoring error type. More specifically, the current study provides original data suggesting that late adolescents exhibited a greater rate of externalizing misattributions (confusing word items the participant generated for words generated by the experimenter) than young adults and this was particularly pronounced for negatively charged items. This finding may have clinical relevance, given that in contrast to young adults, the late adolescent group in the current study displayed a pattern of reality-monitoring errors (i.e.

externalizing misattributions for negative material) that has also been identified among patients suffering with schizophrenia and hallucination-prone adults. On the basis of the current data, it appears that a certain degree of externalizing reality-monitoring confusions for negative material during late adolescence represents a normal aspect of human variation, which attenuates in young adulthood. This tentatively suggests that the pattern of externalizing reality-monitoring misattributions often observed in clinical and non-clinical adult populations experiencing hallucinatory-phenomena may reflect the developmental breakdown of neurobiological and psychological processes normally responsible to sustain the normative elaboration of reality-monitoring from late adolescence towards young adulthood. Thus, from an early prevention standpoint, the period between late adolescence and early adulthood may constitute a key developmental window to psychotherapeutically support young people at increased risk for psychosis sustain resilience against the development of maladaptive reality-monitoring patterns. For instance, psychotherapeutic treatments that focus on helping young people regulate their emotional reactivity may also support their capacity to undertake accurate reality-monitoring judgements, thus attenuating the risk for the emergence of clinical manifestations in adulthood.

Finally it must be noted that the early adolescent group in this study exhibited a higher rate of externalizing errors for positive material (confusing self-generated word items for words generated by the experimenter) compared to the young adult group. This finding suggests that the maturation of the processes sustaining the capacity to monitor the source of positive material continues from early adolescence to young adulthood. Interestingly, an externalizing attribution for self-generated positive material has previously been associated with symptoms of depression (Kinderman and Bentall, 1997). Thus, it is possible the normative elaboration of reality-monitoring processes for positive material during adolescence may confer increased resilience against clinical depression in adulthood.

4.4.2 Associations between reality-monitoring errors and schizotypal trait dimensions

The present analyses suggest that neither the total rate of reality-monitoring errors, nor the rate of externalizing reality-monitoring errors were associated with the cognitive-perceptual of schizotypy in adolescence. Furthermore, the total rare of reality-monitoring errors was not associated with either the interpersonal or the disorganization dimensions of schizotypy.

The lack of associations between positive schizotypy and reality-monitoring was unexpected. Most studies that have examined reality-monitoring for speech in relation to schizotypy, albeit in adult samples, have reported links with positive manifestations (Morrison and Haddock, 1997; Larøi et al., 2004). It must be noted however, that contrary to the current study, previous studies have not used correlational designs to test the associations between reality-monitoring and schizotypal traits. Rather these have reported between-group differences in reality-monitoring performance when comparing those with high scores on measures of positive schizotypy and hallucination-proneness to those with low scores on these measures. Therefore, one possibility is that the link between reality-monitoring and positive schizotypal manifestations can be only observable among individuals that are further along the continuum of psychosis expression. Indeed, the current sample of adolescents and young adults exhibited relatively low scores on the positive dimension of schizotpy. Interestingly, the only study to date that has examined the capacity to differentiate between self- and other-generated material in a CHR sample also did not report associations between reality-monitoring errors and specific symptom dimensions (Johns et al., 2010). Thus, another possibility is that reality-monitoring tasks are not sensitive enough to capture associations between reality-monitoring errors and different symptom dimensions when these are assessed using correlational designs .

The results of the current study also contrast those of the second study of the current thesis that reported links between the interpersonal dimension of schizotypy and total rate of self-monitoring errors for speech (confusions between silently and overtly read words) in a sample of community adolescents (Salaminios et al., 2020b). On the one hand this is surprising given that both reality-monitoring (discriminating between self-generated and externally produced evens) and self-monitoring (discriminating between covert/imagined and overt/enacted events) rely on the capacity distinguish internally-generated from perceptually-based information (Docherty, 2012). On the other hand, both behavioural and neuroimaging data suggest that although interrelated, reality-monitoring and self-monitoring represent separate processes (Stephane, 2019) and the former appears to be relatively easier than the latter (Lagioia et al, 2011; Folley et al., 1983). Therefore, as mentioned above, in contrast to self-monitoring tasks, reality-monitoring paradigms may only capture associations with schizotypal manifestations among those who score at the higher end on measures of schizotypy and are considered to be further along the continuum of psychosis expression.

4.4.3. Limitations and conclusions

The results of the current study should be interpreted in light of certain limitations. First, the data were derived from a relatively small sample and no prior power analyses were conducted to determine the sample size needed to detect significant effects. Thus, it remains possible that further associations between schizotypal personality traits and reality-monitoring errors could have emerged with a larger sample. It must be noted however that to further explore whether the results of the correlation analyses represented true effects as opposed to non-conclusive findings and to reduce the probability of type II errors, Bayesian statistics were obtained and these provided support in favour of the study's findings. Nonetheless, future studies should test the results of the current study with larger samples. Second, the analyses performed were cross-

sectional, and longitudinal investigations are needed to examine causal associations between schizotypy and reality-monitoring. Third, it must be noted that the mixed ANOVAs performed to statistically analyse the effects of the two within-subject factors (i.e. reality-monitoring error type and item valence type) on reality-monitoring performance did not account for individual differences between the participants included in the sample. Therefore, the current study could have benefited by utilizing multilevel modelling statistical approaches to incorporate subject-related random effects to the within-subjects effects, thus account for individual differences in participant's recognition and self-monitoring performance under the different experimental conditions. Furthermore, it must be noted that in the current study the different SPQ dimensions correlated quite highly with each other. Given the clinical focus of its items, it is possible that in the current sample the SPQ may have tapped into overlapping aspects of schizotypal trait expression that cut across its different dimensions (e.g. the level of subjective distress) rather than capturing the unique characteristics of the different dimensions.

Importantly, a number of limitations pertaining to the experimental task selected to test reality-monitoring also need to be highlighted. First, the word items presented by the experimenter may have differed in terms of the semantic and associative memory consequences they generated both within and between different participants and these may have confounded performance in the reality-monitoring test. For instance the presentation of the neutral word "orange" by the experimenter may have activated a number of sensorimotor traces (e.g. the taste of an orange, the act of peeling of an orange, etc.), thus potentially confounding participants' performance when distinguishing in memory on whether it was generated by themselves of the experimenter. In contrast, the presentation of the neural word "cravete" (meaning "necktie") by the experimenter may not have activated strong sensorimotor cues to participants that have never put on a necktie before, thus potentially supporting its identification as other-produced during the reality-monitoring test. In addition, when

generating their own words, participants were free to respond to experimenter-produced items with any word of their choice and no restrictions were imposed on the selection of these, thus a number of confounding factors pertaining to the level of processing involved (e.g. semantic processing), as well as the strategies used to generate and memorise self-produced items may have affected reality-monitoring performance. For instance, a participant who is presented with the neutral word "necktie" by the experimenter and is used to wearing striped neckties may generate the word "stripe". Given the personal relevance of the word "stripe" and the associative strategy involved in its generation the participant may be more likely to accurately identify its source as self-generated rather than other-produced in the reality-monitoring test. Interestingly, previous research indicates that the semantic processing of word items may play a part in the association between externalizing source-monitoring misattributions and hallucination-like experiences (Sugimori et al., 2011). However, the current study did not address the level of processing engaged either during the presentation of experimenter-produced words or during the generation of new words by participants.

Despite these limitations, this study is the first to directly examine reality-monitoring for speech in typically-developing adolescents and the nature of its associations with personality traits pertaining to schizotypy. The present analyses highlight the importance of emotional valence on the level of reality-monitoring confusions across adolescence and young adulthood, as well as the effects of age on reality-monitoring patterns for emotionally charged material. From an early prevention and intervention standpoint, current findings suggest that late adolescence may represent a critical developmental window to support the elaboration of reality-monitoring processes for negatively charged material in order to attenuate psychosis risk in adulthood.

Chapter 5

Prospective effects of schizotypal trait expression and self-monitoring during adolescence on ToM Performance at five-year follow-up

5.1.1. Introduction

As discussed in section 1.5.1 Theory of mind (ToM) – the capacity to make inferences about other people's mental states, such as beliefs, feelings and intentions (Premack and Woodruff, 1978) - represents an integral part of a constellation of mentalizing functions considered to be impaired in schizophrenia (Brüne, 2005; Harrington et al., 2005; Sprong et al., 2007). Given that a number of core psychotic experiences, such as delusions and hallucinations, explicitly entail the misunderstanding of other people's intentions, a large body of empirical research has investigated ToM in individuals suffering with schizophrenia. Indeed, meta-analytic investigations have shown that disruptions in the capacity to understand the inner states underpinning others' behaviours are consistently present among people with long-standing psychotic experiences (Brüne, 2005), as well and in those presenting with first episode psychosis (FEP) (Bora and Pantelis, 2013). Furthermore, evidence suggest that impaired ToM is linked to the level of psychotic symptomatology in those suffering with the illness (Frith, 2004).

Interestingly, studies exploring the associations between specific symptom sub-groups and ToM have generated inconsistent results. For instance, a number of studies have found that poor performance in ToM is primarily associated with negative symptomatology (Corcoran and Frith, 2003; Corcoran et al., 1995; Weijers et al., 2018), while others have reported associations with positive (Sarfati et al., 2000; Frith and Corcoran, 1996; Versmissen et al., 2008) and disorganization symptoms (Sarfati et al., 1999; Mazza et al., 2001). One potential reason that can account for the observed divergence in findings pertains to methodological

differences in the measurement of ToM across different studies. For example, some studies have focused on false belief understanding by using either first- (i.e. understanding what one is thinking of feeling) or second-order ToM tasks (understanding what another person thinks or feels about what a third person in thinking or feeling), while others have focused on the assessment of irony and metaphor perception. In a similar vein, some studies tend to use tasks in which social interactions are presented to participants pictorially (i.e. in the form of cartoon strips), while others utilise task-based measures that involve the verbal or written presentation of social scenarios (Brüne, 2005).

Furthermore, although ToM is a complex and multifaceted process that involves the attribution of both thoughts and feelings in others (Shamay-Tsoory et al., 2007), the majority of experimental tasks used in patient samples tend to only measure singular components of ToM (Sharp et al., 2011), thus do not distinguish between inferences made for cognitive and affective states (Montag et al., 2011). For instance, false belief tasks tend to assess participants' reasoning understanding of differences between the speaker's and listener's knowledge about beliefs. Conversely, ToM tasks that assess the understanding of irony and faux-pas also require an empathic consideration of emotional states in others. This is important, given that the limited number of studies that have used experimental tasks that concurrently assess and distinguish between affective and cognitive ToM dimensions suggest that the former is associated with both positive and negative psychotic symptoms, while the latter appears to relate primarily to positive symptoms (Shamay-Tsoory et al., 2007; Montag et al., 2011).

Overall, variations in the methodological approaches adopted in the measurement of ToM in psychosis are likely to yield differential associations with different symptom dimensions. Nonetheless, ToM performance, across a variety of measures, has been consistently shown to be impaired in people suffering with schizophrenia and appears to relate to the clinical manifestations of the illness.

Critically, ToM difficulties are shown to already be present during the early stages of emerging psychosis, among help-seeking individuals that are at clinical high-risk for psychosis (CHR), prior to transition to transition to the first diagnosable episode of the illness (Chung et al., 2008). Furthermore, data from prospective research suggest that although ToM difficulties in CHR samples may not represent consistent "predictors" of transition to clinical psychosis, they do confer increased vulnerability for the illness among those who are at increased risk (Kim et al., 2011; Piskulic et al., 2016; Debbané, Salaminios, et al., 2016). While contemporary research has examined ToM difficulties during the later stages of psychosis-vulnerability (i.e. in the prodromal stages of the illness), there remains a need to investigate these among non-clinical individuals who report schizotypal trait manifestations, prior to the development of clinically-relevant symptoms. This is important given that poor ToM among individuals who already exhibit clinical or attenuated manifestations of psychosis (i.e. patients suffering with schizophrenia and CHR samples) has consistently been associated with poor clinical and functional outcomes that often remain stable following psychotherapeutic and pharmacological treatment (Armando, Hustebaut and Debbané, 2019). Understanding the links between schizotypal traits and ToM can therefore support the implementation of preventative interventions targeting mentalizing abilities to attenuate psychosis vulnerability and promote better functional outcomes among those who are at trait risk for the illness.

A relatively small amount of cross-sectional research has examined the associations between performance in basic ToM skills and schizotypal trait dimensions in non-clinical samples. Studies in adult samples have reported that high scores on the cognitive-perceptual dimension are associated with reduced ToM scores (Langdon & Coltheart, 1999; Pickup, 2006). In line with these findings, the only study to date that tested the relationship between ToM performance and schizotypal dimensions in a non-clinical adolescent sample also reported an inverse association with the cognitive-perceptual dimension of schizotypy

(Barragan et al., 2011). Overall, ToM research in both adult and adolescent samples from the community indicate that impairments in the capacity to make accurate inferences about other people's mental states relate primarily to cognitive-perceptual manifestations of schizotypy, such as hallucination and delusion-like phenomena (Pickup, 2006). It must be noted however that the schizotypy research described above has only tested basic ToM abilities by using tasks that were predominantly developed for children (thus may yield ceiling effects). Furthermore, these involve either pictorially- or verbally-presented social situations that do not approximate the demands of real-life social situations. At present, no study has investigated the associations between schizotypy and ToM by using age-appropriate tasks with high ecological validity that can adequately address the complex nature of ToM inferences.

Furthermore, despite findings from cross-sectional studies, the nature of the relationship linking the expression of schizotypal personality traits to the development of ToM remains unclear. Developmentally, the maturation and elaboration of ToM progressively continues throughout adolescence, into early adulthood (Choudhury et al., 2006; Dumontheil et al., 2010; Vetter et al., 2013). This is important because schizotypal trait manifestations commonly emerge prior to adulthood, during adolescent development (Gooding et al., 2005). As such, cognitive and interpersonal aberrations arising in the context of schizotypal trait expression during adolescence may disrupt the normative development of ToM, impair the capacity for mental state understanding and further increase the risk for clinical psychosis (Debbané & Barrantes-Vidal, 2014). Indeed, recent evidence indicate that young people who display attenuated psychotic manifestations demonstrate reduced trajectories for common age-related improvements in ToM compared to healthy controls (Davidson et al., 2018). Similarly, the expression of negative schizotypy may in itself impact on the opportunity to establish and maintain close interpersonal relationships with others, which normally sustain the developmental elaboration of ToM (Armando, Hustebaut and Debbané, 2019; Salaminios and

Debbané, 2021). This is in line with the results from study 1 of the current thesis (Chapter 2) which showed that the expression of schizotypal manifestations that impede interpersonal communication within social situations (i.e. social anxiety and disorganized speech) are cross-sectionally linked to self-reported mentalizing difficulties during adolescence (Salaminios et al., 2020a). In a similar vein, data from a retrospective study have found associations between diminished mentalizing capacities in adolescence and reports of social anxiety in childhood (Ballespí et al., 2018).

Notwithstanding evidence from cross-sectional and retrospective studies, there is currently a paucity of research examining prospective associations between schizotypy and ToM during the critical developmental window spanning from adolescence to early adulthood. Understanding the nature of prospective associations linking schizotypal trait manifestations to ToM dysfunction can have important clinical implications for the application of early prevention treatments during the premorbid stages of psychosis expression, prior to the development of clinical symptoms.

It has previously been proposed that ToM difficulties in psychosis may be explained by multiple underlying mechanisms and these may also characterise the symptomatic expression of the illness (Frith, 2004). According to Frith (2004), patients who exhibit predominantly negative symptoms and disorganized thinking, experience a deficit in the representational abilities necessary to attribute mental states, thus display a *hypomentalizing* pattern characterised by reduced use of inner mental states to understand others' behaviours. Conversely, patients who primarily exhibit positive symptoms, such as paranoid delusions or delusions of reference, may possess the basic representational abilities needed to understand mental states, but tend to apply these in a maladaptive way by overattributing intentional mental states to others in the absence of observable data to support their inferences (Frith,

2004). This latter pattern has been referred to in the literature as "hypermentalizing" (Sharp et al., 2011) or "hyper-ToM" (Clemmensen et al., 2014).

Interestingly, most ToM tasks only use dichotomous ("right/wrong") response formats and are reliant on linear scales of overall accuracy (Montag et al., 2011; Poznyak et al., 2019), thus may not be sensitive enough to capture hypermentalizing misattributions. Furthermore, traditional ToM tasks present participants with artificial experimental environments that only assess singular ToM modalities, thus are limited in their ability to assess the complexity of ToM as it unfolds within every day social interactions (Montag et al., 2011).

Critically however, the small number of studies that have utilized task-based measures specifically geared to distinguish between hypomentalizing and hypermentalizing errors in ToM have found associations between positive symptoms and the over-attribution of intentions in others among both chronic psychosis (Montag et al., 2011) and first episode sufferers (Bliksted et al., 2019). In line with studies in clinical samples, research among non-clinical adults has shown that individuals reporting high scores on measures of positive schizotypy and delusion proneness exhibit higher rates of hypermentalizing ToM attributions compared to those with low scores on these measures (Fyfe et al., 2008). Furthermore, it has been shown that children who display a hypermentalizing ToM pattern are significantly more likely to report a history of psychotic-like experiences than non-hypermentalizing children and this effect is particularly pronounced for experiences of paranoid/persecutory ideation (Clemmensen et al., 2014). In addition, when compared with other risk factors, such as family illness, concurrent psychiatric diagnoses, involvement in bullying, gender and socio-economic changes, hypermentalizing has been found to be the only factor that independently accounted for a history of psychotic-like experiences in samples of children from the community (Clemmensen et al., 2016).

Together studies in non-clinical child and adult samples suggest that trait-risk for psychosis is associated to hypermentalizing misattributions and the relationship between the two emerges early in development. Given the comprehensive evidence linking psychosis-risk to aberrant ToM, it is important to prospectively investigate the effects of schizotypal trait dimensions on different forms of ToM dysfunction (hypomentalizing vs hypermentalizing) during the critical period of adolescent development. To date, the majority of ToM research in non-clinical subjects has focused solely on cognitive-perceptual manifestations, which are not as enduring as interpersonal and disorganization features of schizotpy. While studies in child samples have reported that hypermentalizing is associated with psychotic-like experiences, such as delusional ideation and hallucinatory phenomena (Clemmensen et al., 2014), it remains unknown whether the early expression of interpersonal and disorganization features of schizotpy, which are known to confer increased risk for clinical psychosis, also relates to the tendency to make hypermentalizing inferences. Understanding how the early expression of schizotypal trait dimensions during adolescence prospectively relates to specific types of ToM dysfunction may further inform the application of targeted early prevention treatments aiming to sustain mentalizing functioning among young people who are at increased risk.

Another relatively unexplored area of empirical interest involves the potential links between processes that sustain the awareness of the self and others. As discussed in chapter 1, contemporary research suggests that disruptions in thinking abilities pertaining to the understanding of the self (i.e. self-monitoring) and others (i.e. ToM) are both present in schizophrenia and independently contribute to the developmental unfolding of the illness during its premorbid and prodromal stages (Brent and Fonagy, 2014). A key question raised by these findings pertains to whether the observed overlap between impairments in self-monitoring and ToM among people suffering with schizophrenia and high-risk individuals is

indicative a more fundamental disturbance in the recognition of the boundaries between internal and external reality (Murphy et al., 2010).

Indeed, most theoretical accounts seeking to explain the clinical manifestations of psychosis implicitly or explicitly articulate links between self-monitoring and ToM. For example, it has been proposed that auditory hallucinations and delusions of alien control are underpinned by difficulties in monitoring the origin of one's own self-generated inner speech and physical actions, which are instead attributed to an external agents (Allen et al., 2007; Blakemore et al., 2003). Inherent in these conceptualizations is the idea that the externalizing misattribution of one's own self-generated mental events may also lead to a distorted perception and interpretation of the external agent's mental states. For instance, if a malevolent inner voice is experienced as externally-perceived and attributed to external agents, it is likely that their intentions will also be subjectively perceived as threatening and persecutory (i.e. hypermentalizing). In a similar vein, misattributing something one said to another person as something that one only imagined saying (i.e. internalizing self-monitoring misattribution), may prevent them from forming an accurate mental representation of the other person's behavioural response (i.e. hypomentalizing).

The conceptual link between self-monitoring and ToM has been empirically supported by behavioural data. For example, Fisher et al. (2008), examined the relationship between performance in a self-monitoring task (assessing the capacity to differentiate between self-generated and externally-perceived word items) and ToM in a sample of schizophrenia patients (n = 91) and a group of healthy controls (n = 30). In both groups, self-monitoring confusions were significantly and negatively correlated with scores across different measures of ToM performance. In line with behavioural findings, neuroimaging research has consistently reported shared neural underpinnings between the two processes (Murphy et al., 2010). More specifically, functional magnetic resonance imaging (fMRI) data from samples of healthy

adults show relative activation in the medial prefrontal cortex (mPFC) during a wide range of ToM processes (Brunet-Gouet and Decety, 2006), as well as during self-referential cognitive processing, including self-monitoring (Vinogradov et al., 2006; Simons et al., 2006; Murphy et al., 2010).

While behavioural and neuroimaging data support the link between self-monitoring and ToM, developmental research suggests that the two may follow different developmental trajectories from childhood to early adulthood. More specifically, while the elaboration of ToM is shown to continue throughout adolescence and into early adulthood (Dumontheil et al., 2010), self-monitoring appears to be established prior to adolescence, during childhood development (Sussman, 2001; Salaminios et al., 2020b). Therefore, given their neural and behavioural overlap, one interesting possibility is that self-monitoring difficulties in adolescence may be prospectively predictive of ToM dysfunction further along development. This would be in line with contemporary clinical models of psychosis development suggesting that impairments in self-referential processing during the premorbid period of psychosis expression, particularly in the context of high psychometric schizotypy, may contribute to the abnormal explanations of social experience commonly observed during the prodromal and clinical stages of the illness (Brent and Fonagy, 2014; Debbané, Salaminios et al., 2016).

However, no study to date has examined the prospective effects of self-monitoring difficulties or their interactions with schizotypal trait dimensions on ToM dysfunction. Furthermore, despite evidence suggesting cross-sectional associations between self-monitoring and ToM performance, it remains unclear whether specific types of self-monitoring misattributions (i.e. externalizing vs. internalizing) relate to different forms of ToM dysfunction (hypermentalizing vs. hypomentalizing).

The present study will examine ToM performance in a sample of community adolescents and young adults using the Movie for the Assessment of Social Cognition (MASC,

Dziobek et al., 2006) – an ecologically valid and sensitive video-based task that offers the unique advantage of independently assessing the use of aberrant ToM strategies, such as hypermentalizing and hypomentalizing, for both cognitive and affective mental state inferences. The main aims of the study are to test the effects of schizotypal trait dimensions, self-monitoring difficulties and their interactions assessed during adolescence, on ToM performance measured at 5-year-follow-up. To test these effects, all adolescent subjects that participated in study 2 of the current thesis and had completed assessments of schizotypal traits and self-monitoring were re-contacted and invited to complete the MASC after a 5-year interval following their participation in study 2. Due to practical considerations and to reduce participant attrition, schizotypal traits and self-monitoring performance were not assessed at follow-up. Furthermore, it must be noted that participants did not complete the MASC at baseline. As such the current study did not assess how changes in schizotypal trait expression and self-monitoring performance over time account for developmental changes in ToM. Rather, the current study only assessed whether self-reported schizotypal traits and self-monitoring performance assessed during adolescence relate to ToM scores measured 5 years later.

Previous research has reported that non-clinical children and adults scoring high on measures of psychotic-like experiences and delusion proneness exhibit hypermentalizing ToM errors when making inferences about other people's thoughts and intentions (Fyfe et al., 2008; Clemmensen et al., 2014). Thus, it is hypothesised that higher scores on the cognitive-perceptual dimension of schizotypy in adolescence will prospectively relate to increased hypermentalizing errors for cognitions. Because the majority of ToM studies in non-clinical populations have focused on cognitive-perceptual manifestations and no study to date has prospectively examined the relationship between schizotypal dimensions and ToM in adolescence, the current study will also explore whether higher scores in the interpersonal and

disorganization dimensions of schizotypy, will prospectively relate to ToM performance for cognitions and affects. More specifically, in accordance with data from clinical samples (Corcoran and Frith, 2003; Corcoran et al., 1995, Montag et al., 2011), and on the basis of findings from study 1 of the current thesis (Salaminios et al., 2020a), it is expected that the interpersonal and disorganization dimensions of schizotypy will prospectively relate to ToM errors for mental state inferences involving both cognitions and affects.

Regarding the investigation of prospective associations between processes that sustain self- and other-awareness, the current study seeks to explore whether difficulties in cognitive abilities pertaining to self-monitoring during adolescence will prospectively relate to specific types of ToM dysfunction 5 years later. Following previous cross-sectional data linking self-monitoring to ToM (Fisher et al., 2008), it is hypothesised that the rate of externalizing self-monitoring misattributions (i.e. confusing imagined mental events as real) will prospectively account for the rate of hypermentalizing errors for cognitions. Furthermore, it is hypothesised that the rate of internalizing self-monitoring (i.e. confusing real events as imagined) will prospectively account for the rate of hypomentalizing errors for cognitions.

Finally, the current study aims to provide a preliminary examination on whether trait risk for psychosis and self-monitoring difficulties interact during adolescence to prospectively account for ToM difficulties in terms of hypermentalizing and hypomentalizing errors for cognitions and affects. Given clinical conceptualizations of psychosis development (Brent and Fonagy, 2014; Debbané, Salaminios et al., 2016) and findings from previous studies suggesting that both schizotypal trait expression and self-monitoring difficulties relate to ToM performance (Fisher et al., 2008; Pickup; 2006), it is expected that high scores on schizotypal dimensions will account for increased hypermentalizing and hypomentalizing ToM scores at high levels of self-monitoring errors.

5.2. Methods

5.2.1. Participants and procedure

All subjects that participated in study 2 of the current thesis (see chapter 3, Salaminios et al., 2020b) and had completed assessments of schizotypy and self-monitoring were re-contacted and invited to complete an assessment of ToM abilities after a 5-year interval following the original testing. No prior power calculations were performed to determine sample size. The original sample from study 2 included 121 young people (12-18 years of age) who were recruited through written advertisements in public schools in the city of Geneva, Switzerland, were fluent in French and had a score of 7 or above in the Block Design subtest of the Wechsler Intelligence Scale (Wechsler, 1955).

From these, fifty-seven (n = 57) agreed to take part in the follow-up assessment, while sixty-four (n = 64) either declined participation or did not respond to the invitation. Participants that completed the follow-up assessment more than 5 years after the initial assessment were excluded from the final analyses (n = 7). The final sample consisted of 50 community adolescents and young adults (24 females, 26 males). None of the participants suffered from past/present psychiatric, neurological or neurogenetic disorders at the time of testing. Written informed consent was obtained from all participants and legal guardians of those under 18 years of age. Ethical approval for the study was granted by the Swiss Ethics Committee on research involving humans (number: 2018-00251).

Scores on a self-report measure of schizotypal trait dimensions and on a self-monitoring task, assessed as part of study 2 of the thesis, were used as the baseline (T1) data for the current study ($M_{age\ at\ T1}$ = 15.20; SD = 1.76; range = 12.01-17.95). At the second assessment (T2), which took place after a 5-year interval from initial testing ($M_{interval}$ = 5.01; SD = 0.11), participants were asked to complete a video-based task assessing the capacity to make inferences about other people's mental states ($M_{age\ at\ T2}$ = 20.21; SD = 1.75; range 17.01-23.10).

It must be noted that because of practical considerations and to reduce participant attrition, subjects were not asked to complete assessment of schizotypy and self-monitoring at follow-up. Furthermore, participants did not complete the MASC at baseline. As such, the measures administered differed between baseline and follow-up.

5.2.2. Measures

5.2.2.1. Baseline (T1) measures

The Schizotypal Personality Questionnaire (SPQ, French version, (Dumas et al., 2000) was used at T1 to measure schizotypal traits, yielding three dimension scores and nine subscale scores: cognitive–perceptual (unusual perceptual experiences, ideas of reference, suspiciousness, odd beliefs or magical thinking), interpersonal (social anxiety, constricted affect, lack of close friends), and disorganization (odd speech, odd behaviour). The SPQ has been validated for French-speaking adolescents (Badoud et al., 2011).

Self-monitoring at T1 was investigated using a speech-monitoring task (Debbané et al., 2010) consisting of two parts: a reading procedure, followed by a recognition and self-monitoring test. In the first phase participants were required to read, either aloud or silently, a series of common words (low cognitive effort) or non-words (high cognitive effort) presented on a computer screen.

During the reading procedure, six blocks (three silent, three aloud) of eight items (four words, four non-words) were randomly presented. Each condition contained 12 items, for a total of 48 items (12 word/overt items, 12 non-word/overt items, 12 words/silent items and 12 non-word/silent items). After a 15-minute filler task, a recognition sheet was handed out, containing 72 items (48 original items,12 new word and 12 new non-word items). In the test phase, participants had to indicate if each item from the list appeared in the first part of the task and attribute recognized items to a reading condition, silent or aloud (self-monitoring test). The

task yields two types of self-monitoring errors: externalizing errors (attributing silently-read items as overtly-read), and internalizing errors (attributing overtly-read items as silently-read).

The percentage of externalizing misattributions was calculated by dividing the total score for items read silently but identified as read aloud in the monitoring test, out of the total number of correctly recognized silent items. The percentage of internalizing misattributions was calculated by dividing the total score of items read overtly but identified as read silently in the monitoring test, out of the total number of correctly recognized overt items.

5.2.2.2. Follow-up (T2) measures

The Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006) is a computerized video-based task for the assessment of ToM inferences that approximate the demands of everyday life. The task adopts traditional ToM constructs, such as first and second-order false belief, faux pas, metaphor and sarcasm for both visual and auditory input and operationalizes these through a short movie approximating real-life everyday interactions. The current study employed the French version of the task during the follow-up assessment.

Participants were asked to watch a 15-min movie featuring four characters interacting during a dinner party. The movie plot involves the interaction of these characters getting together for an evening of cooking, dining, and playing a board game. Each character encounters different situations throughout the course of the film that elicit emotions and mental states such as anger, affection, appreciation, jealousy, fear, ambition, embracement and disgust. The relationships between the different characters vary in terms of intimacy (i.e. friends or strangers), thus providing different social reference systems on which mental state inferences have to be made.

The movie stops at different moments during the plot and participants are asked a multiple choice question about the characters' mental states during the interaction (i.e. what the

characters are feeling or thinking at the very moment the film is stopped). The questions mostly refer to complex mental states and allow the identification of subtle mentalizing difficulties. Questions and multiple-choice options are read aloud by the experimenter and silently by the participant. The task comprises of 45 multiple choice test questions, resulting in a maximum total score of 45. The task further provides scores for the following mental state modalities: (1) cognitions (e.g. attribution of thoughts and intentions; e.g. "Why did Anna say that?", "What is Ben thinking?"; 27 items) and (2) emotions (e.g. attribution of anger and guilt; e.g. "What is Mary feeling?"; 18 items). In addition, 6 control questions requiring non-social inferences are asked to control for correct understanding of the task (see appendix 6 & 7 for task instructions and examples).

For each question, four answer options are provided, each option characteristic of a particular mentalizing pattern: (a) Correct mentalizing, referring to balanced inference of others' mental states; (b) Hypermentalizing, indicating an excessive allocation of mental states in others; (c) Hypomentalizing, referring to reduced attribution of others' mental states; and (d) Lack of mentalizing, representing a complete absence of mentalizing. For the purposes of the current study, MASC error scores for cognitions and emotions were analysed separately in terms of the number of Hypermentalizing and Hypomentalizing errors.

5.2.3. Statistical analyses

5.2.3.1. Preliminary analyses

The main statistical analyses were performed in SPSS version 24. Prior to testing the main study hypotheses, preliminary analyses were run to determine means, standard deviations and ranges for the main study variables. In order to examine whether the results of the current study were significantly affected by the rate of participant attrition at follow-up, Mann-Whitney U and independent samples T-tests were run to respectively examine baseline differences in SPQ

and self-monitoring scores between participants that were included in the main analyses (n = 50) and those that were lost to follow-up (n = 64).

Next, Spearman's correlation coefficients were calculated to examine baseline cross-sectional associations between SPQ scales and self-monitoring error scores (externalizing and internalizing) at T1.

Given that recent findings have shown that age and gender are cross-sectionally associated with ToM performance (Poznyak et al., 2019), their bivariate relations with MASC scores at T2 were also assessed to determine whether they represent confounding factors needing to be accounted for in the main analyses. Spearman's correlations coefficients were used to assess the relationship between age at T2 and MASC hypermentalizing and hypomentalizing errors for cognitions and emotions. Furthermore, Mann-Whitney U tests were used to examine differences between male and female participants on hypermentalizing and hypomentalizing inferences for cognitions and emotions as measured by the MASC.

5.2.3.2. Main Analyses

Hierarchical linear regression models were computed to examine the predictive value of SPQ subscales (cognitive-perceptual, interpersonal, disorganized), self-monitoring error scores (externalizing and internalizing) and their interactions at T1 (independent variables), on MASC hypermentalizing and hypomentalizing scores for cognitions and emotions at T2 (dependent variables). In each case, a regression analysis was computed with SPQ scales entered at the first step (indicating their unique contribution); and externalizing and internalizing self-monitoring errors entered at the second step (revealing their effects over-and-above the SPQ scales). The interactions between SPQ scales and the two separate types of self-monitoring errors (SPQ cognitive-perceptual × externalizing errors; SPQ interpersonal × externalizing

errors; SPQ disorganised × externalizing errors; SPQ cognitive-perceptual × internalizing errors; SPQ interpersonal × internalizing errors; SPQ disorganised × internalizing errors) were entered at the third step to examine their contribution over and above the main effects.

5.2.3.3. Bayesian analyses of regression

In order to test the results of the main analyses exploring the effects of SPQ scales and self-monitoring errors at T1 on MASC performance at T2, JASP version 0.9.2 was used to obtain simple Bayesian statistics of the regression models ran (Van den Bergh et al., 2021;). In contrast to standard hypothesis testing based on P values, Bayesian statistics offer a direct measure of the strength of the evidence both in favour and against the study hypotheses, thus supporting the interpretation of results. More specifically, rather than basing inferences on a single model, Bayesian regression analysis retains all possible models for inference and quantifies the importance of individual predictors by providing a Bayes factor ($BF_{inclusion}$), which signals the factor by which the prior probabilities of including a predictor in a model (calculated by dividing the number of models that contain the predictor from the number of all possible models) increase or decrease after observing the data (i.e. after averaging the predictive odds in favour of the models that contain the individual predictor). $BF_{inclusion}$ values below 1.00 show the factor by which the data have decreased the prior odds for including a predictor in a model, while values above 1.00 show the factor by which the prior odds for including a predictor in a model have increased based on the data.

5.3. Results

5.3.1. Descriptive results

Table 5.1. presents the descriptive results for the major variables included in the analyses. No significant differences were identified between those that completed the follow-up assessment and those that were lost to follow-up in terms of the cognitive-perceptual (U = 1630.50; p = 0.20); interpersonal (U = 1877.0; p = 0.98); and disorganized (U = 1391.50; p = 0.08) scales

of the SPQ, or in terms of externalizing ($t = -1.53$, $p = 0.22$) and internalizing ($t = -2.15$, $p = 0.97$) self-monitoring errors at T1.

	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
MASC (T2)						
Hypermentalizing for cognitions	4.34	1.65	1.00	8.00	0.51	-0.06
Hypermentalizing for emotions	2.12	1.56	0.00	7.00	0.70	0.60
Hypomentalizing for cognitions	2.54	1.12	0.00	5.00	0.09	-0.75
Hypomentalizing for emotions	2.62	1.63	0.00	7.00	0.65	0.54
Self-monitoring Test (T1)						
Self-monitoring errors Total (%)	29.10	19.44	0.00	55.21	-0.30	0.44
Internalizing errors Total (%)	30.89	15.33	0.00	58.33	-0.01	-0.57
Externalizing errors Total (%)	27.31	14.21	0.00	66.67	0.32	0.12
Schizotypal Traits (SPQ, T1)						
SPQ Cognitive-perceptual dimension	10.08	7.73	0.00	33.00	0.97	0.37
SPQ Interpersonal dimension	6.88	4.59	0.00	18.00	0.53	-0.65
SPQ Disorganization dimension	6.62	4.04	0.00	14.00	0.10	-1.08

Table 5.1. Means, standard deviations and ranges for each variable in the final sample

Abbreviations: SPQ, Schizotypal Personality Questionnaire

Kurtosis and skewness of the SPQ scale distributions for the final sample were all within the acceptable range (i.e. all values between -2.00 and 2.00) and no outliers were identified Furthermore, the distribution of SPQ scores from the final sample suggest that the proportion of participants reporting elevated schizotypal traits (1.5 standard deviations above the mean) was comparable across the different SPQ scales (10 % for the interpersonal [n = 5]; 10% for the cognitive-perceptual [n = 5] and 8% for the disorganization [n = 4] SPQ dimensions.

5.3.2. Relationships between schizotypal trait dimensions and self-monitoring errors at baseline

Table 5.2. presents correlations between SPQ scales and self-monitoring errors (externalizing; internalizing; total) at T1 for participants that were included in the main analyses (n = 50). None of the SPQ scales was significantly associated with self-monitoring errors.

		1	2	3	4	5	6
1.	Self-monitoring Errors Total (%)	-	0.73**	0.76**	-0.26	0.16	-0.21
2.	Externalizing Errors (%)		-	0.19	-0.01	0.12	-0.16
3.	Internalizing Errors (%)			-	-0.24	0.11	-0.17
4.	SPQ Cognitive-Perceptual				-	0.61**	0.69**
5.	SPQ Interpersonal					-	0.40**
6.	SPQ Disorganization						-

Note: **$p < 0.01$

Table 5.2. Correlations between self-monitoring error percentages and SPQ scales at T1
Abbreviations: SPQ, Schizotypal Personality Questionnaire

5.3.3. Relationships between age, gender and MASC hypermentalizing and hypomentalizing errors for cognitions and emotions

Table 5.3. presents the Spearman's correlations between age and MASC error scores at T2. Age was not associated with ToM performance in terms of hypermentalizing and hypomentalizing scores for cognitions or emotions.

		1	2	3	4	5
1.	Age at T2	-	0.07	0.05	0.02	0.16
2.	MASC Hypermentalizing for cognitions		-	0.01	0.35*	-0.04
3.	MASC Hypomentalizing for cognitions			-	-0.10	0.38**
4.	MASC Hypermentalizing for emotions				-	-0.23
5.	MASC Hypomentalizing for emotions					-

Note: **$p < 0.01$; *$p < 0.05$
Table 5.3. Correlations between age and MASC error scores at T2

Furthermore, results of Mann-Whitney tests did not reveal significant gender differences in MASC performance between male and female participants in terms of hypermentalizing errors for cognitions ($U = 278.50$, $p = 0.51$); hypomentalizing errors for cognitions ($U = 296.0$, $p = 0.75$); hypermentalizing errors for emotions ($t = 296.50$, $p = 0.76$); and hypomentalizing errors for emotions ($t = 285.50$, $p = 0.60$).

5.3.4. Longitudinal effects of baseline schizotypal trait dimensions and self-monitoring errors on hypermentalizing and hypomentalizing ToM inferences for cognitions and emotions at follow-up

Because no significant effects of age and gender on ToM performance were identified, these were not entered as control variables in the following analyses. In terms of hypermentalizing errors for cognitions, results of the hierarchical regression analysis indicate that higher scores in the disorganized SPQ scale at T1 accounted for lower MASC hypermentalizing error scores for cognitions at T2 ($\beta = -0.44$, $t = -2.63$, $p < 0.05$). Furthermore, higher rates of externalizing self-monitoring errors at T1 accounted for higher hypermentalizing error scores for cognitions at T2 ($\beta = 0.37$, $t = 2.68$, $p < 0.05$). Although the cognitive-perceptual and interpersonal SPQ scales at T1 did not independently account for MASC hypermentalizing errors for cognitions at T2, the SPQ interpersonal × externalizing self-monitoring errors interaction did ($\beta = 0.35$, $t = 2.43$, $p < 0.05$) (Table 5.4.). Simple slope analyses of the interaction revealed that higher scores on the interpersonal SPQ scale at baseline significantly predicted higher T2 MASC hypermentalizing errors for cognitions, at high levels of externalizing self-monitoring errors ($\beta = 0.49$, $t = 2.04$, $p < 0.05$) but not at low ($\beta = -0.29$, $t = -1.28$, $p = 0.21$) or moderate levels ($\beta = 0.10$, $t = 0.58$, $p < 0.57$) (Figure 5.1.).
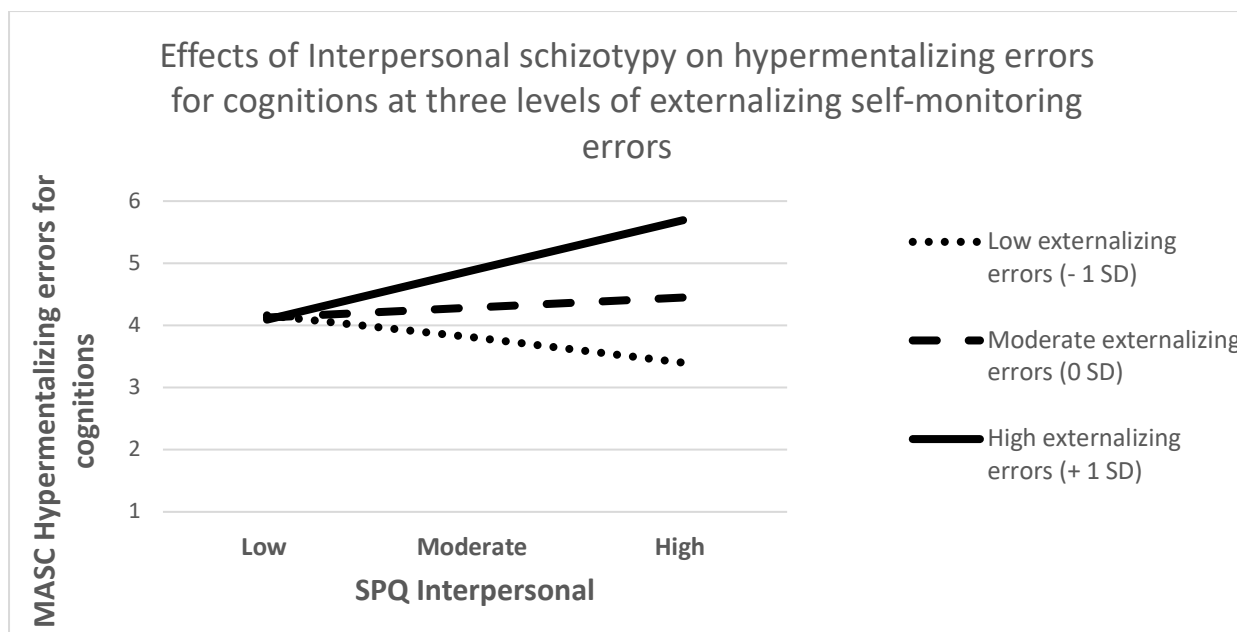
Figure. 5.1. Prospective effects of interpersonal schizotypy on T2 MASC hypermentalizing errors for cognitions at three levels of externalizing self-monitoring errors

Dependent variables	Coefficients			
Independent variables	**B**	**t**	**P**	**R²**
Hypermentalizing errors of cognitions				
Step 1				0.11
SPQ Cognitive perceptual	0.25	1.25	0.22	
SPQ Interpersonal	0.13	0.77	0.44	
SPQ Disorganised	-0.39	-2.19	<0.05	
Step 2				0.22
Externalizing errors	0.36	2.52	<0.05	
Internalizing errors	-0.07	-0.49	0.63	
Step 3				0.44
SPQ cognitive-perceptual × externalizing errors	-0.01	-0.08	0.94	
SPQ interpersonal × externalizing errors	0.35	2.43	<0.05	
SPQ disorganised × externalizing errors	-0.29	-1.73	0.09	
SPQ cognitive-perceptual × internalizing errors	-0.07	-0.32	0.75	
SPQ interpersonal × internalizing errors	-0.32	-1.70	0.09	

SPQ disorganised × internalizing errors	-0.01	-0.04	0.97

Table 5.4. Hierarchical regression of prospective effects of SPQ scales, self-monitoring errors and their interactions at T1 on MASC Hypermentalizing errors for cognitions at T2
Abbreviations: SPQ, Schizotypal Personality Questionnaire

Regarding hypomentalizing errors for cognitions, results of the regression model suggest that higher scores in internalizing self-monitoring errors at T1 accounted for higher MASC hypomentalizing error scores for cognitions at T2 ($\beta = 0.52$, $t = 3.27$, $p < 0.01$). SPQ scales or their interactions with self-monitoring errors at T1 did not significantly account for MASC hypomentalizing error scores for cognitions at T2. Finally, in terms of ToM performance for emotions, SPQ scales, self-monitoring error scores and their interactions at T1 did not significantly account for either hypermentalizing or hypomentalizing error scores atT2.

5.3.5. Bayesian analyses

To test the results of the main analyses presented in section 5.3.4, a series of Bayesian multi-model linear regressions were computed to assess the predictive relevance of SPQ scales and self-monitoring errors at T1 on MASC hypermentalizing and hypomentalizing errors for cognitions and emotions at T2.

First, Bayesian statistics confirmed the findings of the main regression analyses by showing that the data increased the odds of including the disorganization scale of the SPQ ($BF_{inclusion} = 1.23$) and the rate of externalizing self-monitoring errors ($BF_{inclusion} = 5.45$) as predictors of MASC hypermentalizing errors for cognitions. It must be noted however that the BF for including the disorganization scale was very close to 1, suggesting that we should be cautious when drawing conclusions about its predictive effect. Furthermore, the odds of including the SPQ interpersonal × externalizing self-monitoring errors interaction as predictors of MASC hypermentalizing errors for cognitions also increased ($BF_{inclusion} = 1.74$). In contrast,

after observing the data the odds of including the cognitive-perceptual ($BF_{inclusion}$ = 0.55) and interpersonal ($BF_{inclusion}$ = 0.57) scales of the SPQ, as well as the rate internalizing self-monitoring errors ($BF_{inclusion}$ = 0.30) as predictors of MASC hypermentalizing errors for cognitions were reduced.

Furthermore, in line with the findings of the main analyses, Bayesian regressions showed that the data only increased the odds of including the rate of internalizing self-monitoring errors at T1 as a predictor of hypomentalizing errors for cognitions at T2 ($BF_{inclusion}$ = 3.32), while the inclusion odds for the rate of externalizing self-monitoring errors ($BF_{inclusion}$ < 0.59) and SPQ scales decreased ($BF_{inclusion}$ < 0.56).

Regarding hypermentalizing errors for emotions, in accordance with the results of the main analyses, Bayesian regression analyses showed that the data decreased the odds of including externalizing ($BF_{inclusion}$ = 0.64) and internalizing ($BF_{inclusion}$ = 0.33) self-monitoring errors, as well as cognitive-perceptual ($BF_{inclusion}$ = 0.36) and disorganized ($BF_{inclusion}$ = 0.39) SPQ scales as predictors of MASC hypermentalizing errors for emotions. It must be noted that although the data also reduced the odds of including the interpersonal scale of the SPQ as a predictor of MASC hypermentalizing errors for emotions ($BF_{inclusion}$ = 0.94), this was a very small decrease (i.e. $BF_{inclusion}$ value close to 1), thus the possibility that an association between the two could have emerged with a larger sample cannot be excluded.

Finally, regarding hypomentalizing for emotions, results of Bayesian regressions confirmed the findings of the main analyses by showing that the data decreased the odds of including externalizing ($BF_{inclusion}$ = 0.23) and internalizing self-monitoring errors ($BF_{inclusion}$ = 0.48), as well as SPQ scales ($BF_{inclusion}$ = 0.25) as predictors of MASC hypomentalizing errors for emotions.

5.4. Discussion

The current study assessed the effects of self-monitoring difficulties and schizotypal trait expression measured during adolescence, on ToM performance at five-year follow-up. The task employed in the present study resembles the demands of every day social cognition and yields two types of ToM error scores (hypermentalizing and hypomentalizing) for two mental state attribution modalities (cognitions and affects).

Results of the current study suggest that self-monitoring difficulties during adolescence were prospectively associated to ToM difficulties pertaining to cognitions. More specifically, externalizing and internalizing self-monitoring errors respectively accounted for increased hypermentalizing and hypomentalizing ToM scores for cognitions at 5-year follow-up. Furthermore, the disorganization dimension of schizotypy prospectively accounted for decreased hypermentalizing error scores for cognitions. Although no main effects of cognitive-perceptual and interpersonal schizotypal dimensions on ToM performance were identified, a significant interaction effect emerged between the interpersonal dimension of schizotypy and externalizing self-monitoring errors, together accounting for hypermentalizing error scores for cognitions. Further analyses of the interaction effect showed that the interpersonal dimension of the SPQ prospectively accounted for increased hypermentalizing errors for cognitions at high levels of externalizing self-monitoring errors, but not at low or moderate levels. Schizotypal personality dimensions, self-monitoring difficulties or their interactions at baseline did not prospectively account for ToM performance for emotions in terms of hypermentalizing and hypomentalizing errors.

5.4.1. Prospective effects of schizotypal trait dimensions on ToM performance

First, contrary to the hypotheses, no prospective associations were found in the current sample between the cognitive-perceptual dimension of schizotypy and ToM performance in terms of

hypermentalizing and hypomentalizing errors for either cognitions or affects. This is in contrast to previous cross-sectional data from non-clinical adolescent and adult samples showing inverse correlations between positive schizotypy and ToM performance (Pickup, 2006; Barragan et al., 2011), as well as poorer ToM scores among people displaying high scores on cognitive-perceptual manifestations compared to those with low scores (Langdon and Coltheart, 1999). Furthermore, the current findings contrast those of previous studies showing that psychotic-like experiences and delusion proneness in children and adults from the community are linked to a hypermentalizing ToM style (Fyfe et al., 2008; Clemmensen et al., 2014).

The divergence between the current findings and those of previous research can first be attributed to the different types of experimental tasks employed to assess ToM. While the current study used a ToM task in which common social interactions are presented to participants in a video-based form, previous studies in schizotypy have employed ToM tasks that entail either the verbal (Pickup, 2006; Barragan et al., 2011), or pictorial presentation of social interactions (Langton and Coltheart, 1999). Undertaking mental state inferences on the basis of video-recorded social situations, such as the one used in the MASC, enables participants to assess a number of contextual and physical cues as they naturalistically unfold during the characters' interactions (i.e. prosodic information, postural movements, eye contact) (Sharp et al., 2013). Conversely, verbally- and pictorially-presented social interactions provide fewer contextual and physical information, thus require participants to imaginatively envision these in the social situations upon which their mental state inferences will be based. As such, one possibility is that among individuals who experience cognitive-perceptual aberrations, the absence of contextual cues and the additional demands imposed by these tasks in terms of envisioning these, may increase the propensity to overattribute the characters' mental states. This can directly manifest as hypermentalizing in narrative- and picture-based tasks

specifically geared for its assessment (i.e. Clemmensen et al., 2014), or as reduced ToM in tasks that only use dichotomous (right/wrong) response formats (i.e. Pickup, 2006; Barragan et al., 2011). Future studies should test the findings of the current study by examining the cross-sectional associations between schizotypal associations and ToM using ecologically valid video-based tasks, such as the MASC.

Furthermore, the divergence between the present findings and those from previous studies, in terms of the associations between cognitive-perceptual manifestations and ToM, can be attributed to the prospective nature of the analyses performed. As mentioned earlier, previous studies reporting associations between cognitive-perceptual manifestations and ToM difficulties have only done so cross-sectionally and this is the first prospective investigation of their relation. Given the lack of predictive effects of positive schizotypy on ToM performance in the current study and previous findings reporting cross-sectional associations between the two, another possibility is that cognitive-perceptual manifestations in non-clinical individuals may represent outcomes, rather that predictors of ToM dysfunction (Langton and Coltheart, 1999). Similarly, although previous studies have reported that patients with schizophrenia display poor ToM performance when inferring the emotional states of others (Shamay-Tsoory et al., 2007; Montag et al., 2011), no associations were found in the current study between schizotypal dimensions and ToM difficulties in inferring affects, suggesting that the latter may constitute outcomes of transition to clinical illness. It must be noted however that ToM performance at baseline was not measured in the current study. As such, it was not possible to account for the effects of baseline ToM skills when examining the prospective effects of schizotypal traits during adolescence on ToM performance at follow-up. Similarly, schizotypal traits where not measured at follow-up, therefore it was not possible to examine cross-sectional associations between different schizotypal trait dimensions and ToM performance, or whether changes in schizotypal trait expression over time relate to developmental changes in ToM.

Surprisingly, the present findings indicate that self-reported schizotypal manifestations pertaining to cognitive disorganization (i.e. odd speech, odd behaviours) were prospectively linked to lower scores in hypermentalizing errors for cognitions at follow-up. It must be noted however, that the results of Bayesian analyses did not provide conclusive evidence regarding this effect, suggesting that future studies should test the association between cognitive disorganization and hypermenizing for cognitions using larger samples. Nonetheless, given that the current study involved a non-clinical sample, it remains possible that increased awareness and insight about their thinking difficulties, may have enabled those scoring higher on the disorganization dimension of schizotypy to adopt a more conservative style when making mental state attributions about other people's cognitive states, thus displaying a reduced number of hypermentalizing attributions. This would be in line with previous research suggesting bidirectional associations between ToM and insight in people suffering with schizophrenia (Ng et al., 2015).

5.4.2. Prospective effects of self-monitoring difficulties on ToM performance

In accordance with our hypothesis, results of the current study showed that self-monitoring difficulties during adolescence longitudinally predicted ToM performance for cognitions at 5-year follow-up. This resonates with behavioural data suggesting that self-monitoring confusions are cross-sectionally linked to ToM performance in both clinical psychosis and healthy adult samples (Fisher et al., 2008). Current findings are also consistent with functional neuroimaging data showing that self-monitoring and ToM processes rely on common neural substrates (Brunet-Gouet and Decety, 2006; Murphy et al., 2010). The present study adds original findings to this literature, showing that specific types of self-monitoring misattributions are prospectively linked to different forms of ToM dysfunction.

First, current findings indicate that externalizing self-monitoring misattributions in adolescence predicted hypermentalizing ToM errors for cognitions at 5-year follow-up. From a conceptual standpoint this is interesting as both externalizing self-monitoring misattributions and hypermentalizing involve a form of attributional style in which inner imaginative states impinge on the perception of external reality (i.e. what one thinks inside becomes real outside). Within this context, the observed relation between the two may be indicative of a more general tendency to experience one's own imaginative states, pertaining to both the self and others, as equivalent to reality (Fonagy et al., 2002). In the case of the self, this involves the cognitive misattribution of self-generated imagined events (i.e. inner speech) as overtly enacted and real, while in the case of others, it entails the over-attribution of inner states to other people's behaviours. Interestingly, externalizing self-monitoring and hypermentalizing ToM patterns have both been identified across the continuum of psychosis expression. Thus, it can be hypothesised that a form of cognitive "psychic equivalence" (Fonagy and Target, 1996), in which one's mental events in relation to the self and others are concretely experienced as isomorphic to reality, may underpin the reality distortions characteristic of psychotic phenomena (Debbané, Salaminios et al., 2016; Salaminios and Debbané, 2021).

In addition, results of the current study indicate that internalizing self-monitoring misattributions in adolescence prospectively predicted the level of hypomentalizing ToM for cognitions. Indeed, internalizing self-monitoring and hypomentalizing ToM patterns may conceptually overlap, as they both reflect the cognitive tendency to decouple mental events from the perceptually-based domain of external reality. In the case of self-monitoring, this entails the cognitive misattribution of self-generated overt actions as merely imagined, while in the case of hypomentalizing it presents as an inhibition to imaginatively attribute intentional inner states to other people's overt behaviours. Within this context, contrary to the association between externalizing self-monitoring and hypermentalizing, which may be indicative of a

tendency to experience mental events as "too real", the overlap between internalizing-self monitoring and hypomentalizing may reflect the tendency to undermine the significance of inner states on external reality (Fonagy and Target, 1996).

Importantly, individuals who exhibit psychotic phenomena are often shown to concurrently display both hypo- and hyper-mentalizing ToM patterns (Langdon and Brock, 2008). Similarly, it has been shown that adolescents who are at risk for psychosis exhibit increased rates of both externalizing and internalizing self-monitoring misattributions compared to controls (Debbané et al., 2010). As such, future studies can attempt to further elucidate potential affective and interpersonal stress factors that may trigger the use of different attributional styles among the same individuals. This may be clinically-relevant to further inform the development of targeted and personalized preventative interventions during adolescence.

5.4.3. Prospective interaction effects of schizotypal trait dimensions and self-monitoring difficulties on ToM performance

Although current findings suggest that cognitive-perceptual and interpersonal schizotypal dimensions in adolescence did not independently predict ToM performance at 5-year follow-up, a significant interaction effect emerged between the interpersonal dimension of schizotypy and externalizing self-monitoring errors, which accounted for the level of hypermentalizing attributions for cognitions. More specifically, scores on the interpersonal dimension of schizotypy predicted hypermentalizing attributions for cognitions at high levels of externalizing self-monitoring errors, but not at low or moderate levels. This finding has both methodological and clinical implications.

Contrary to clinical psychosis, in which negative symptoms have consistently been associated with ToM performance (Corcoran and Frith, 2003; Corcoran et al., 1995), previous studies in non-clinical adult and adolescent samples have not reported associations between

negative schizotypy and ToM (Pickup, 2006; Barragan et al., 2006). The results of the current study suggest that the observed divergence in findings pertaining to the relationship between negative manifestations and ToM among clinical psychosis and non-clinical samples could be attributed to differences in the level of externalizing self-monitoring misattributions exhibited by these groups. Indeed, studies have consistently shown that patients with schizophrenia demonstrate significantly higher rates of externalizing errors towards recalling imagined stimuli as overtly enacted compared to non-clinical controls (Docherty, 2012; Franck et al., 2000; Henquet et al., 2005). Thus, an increased tendency to make externalizing self-monitoring misattributions may have also accounted for the associations between negative manifestations and ToM reported in clinical samples (Corcoran and Frith, 2003; Corcoran et al., 1995; Weijers et al., 2017). Accordingly, the lack of associations between negative schizotypy and ToM in studies of healthy adults (Pickup, 2006) and adolescents (Barragan et al., 2006) could be potentially attributed to low or moderate levels of externalizing-self monitoring errors in these samples. Given that this is the first investigation of the role of self-monitoring on the relationship between schizotpy and ToM, current findings need be interpreted with caution. Future studies should also test the effects of self-monitoring difficulties on the relation between negative symptoms and ToM in clinical samples.

From a clinical standpoint, current findings suggest that the effects of trait risk for psychosis on the development of aberrant ToM patterns may critically depend upon the level of self-monitoring difficulties during adolescence. Young people who withdraw from social contact in the context of negative schizotypy may limit their opportunities to use close interpersonal relationships to elaboration of their ToM abilities (Armando, Hustebaut and Debbané, 2019). Indeed, results of the first study of this thesis indicate that interpersonal features of trait schizotypy are linked with self-reported mentalizing problems during adolescence (Salaminios et al., 2020a). The current data add to these findings, suggesting that

the expression of interpersonal schizotypy in the context of externalizing self-monitoring confusions can be predictive of the future tendency to overattribute thoughts and intentions to others. This is important given that the particular ToM pattern that has been specifically linked to clinical psychosis (Ciaramidaro et al., 2015). From a dynamic developmental standpoint, it can be hypothesised that in the absence of adaptive interpersonal contact, which normally enhances the development of ToM (Fonagy et al., 2002), young people who display a tendency to confuse self-generated inner thoughts as overtly enacted and real, may extend this pattern, further along development, to their understanding of other people's mental states (i.e. overattribute inner states to others' overt behaviours). The latter may reflect a maladaptive coping strategy used by young people with a compromised sense of self as a means to organize and regulate the increasing complexity of their social interactions. Alternatively or in conjunction, the tendency to confuse self-generated mental events as real, may interfere with young people's ability to function within interpersonal situations, thus increase the tendency to withdraw from social situations and disrupt their developmental capacity to form accurate representations of other people's mental states. Within this developmental framework, current findings suggest that trait risk for psychosis, as represented by the expression of interpersonal schizotypal manifestations, may relate to the development of compensatory aberrant ToM strategies (i.e. hypermentalizing for cognitions), rather than to representational deficits in the capacity to attribute mental states in others (i.e. hypomentalizing).

It has previously been argued that impairments in self-referential processing during the premorbid period of psychosis expression may contribute to the aberrant explanations of social experience commonly observed during the prodromal and clinical stages of the illness (Brent and Fonagy, 2014). Within this context, current findings suggest that self-monitoring difficulties may represent important early prevention treatment targets to sustain resilience against the development of aberrant ToM patterns (i.e. hypermentalizing) and attenuate trait-

risk for psychosis in youths presenting with negative schizotypal features, prior to the development of clinical symptoms. Metacognition- and mentalization-based treatments (Lysaker et al., 2020; Debbané et al., 2016; Salaminios and Debbané, 2021), which specifically aim to help individuals recognise, think about and reflect on their thoughts, as well as synthesise them to develop a more coherent sense of self and personal agency, may be applied preventatively to attenuate disturbances of social understanding in young people during the premorbid period of psychosis expression.

5.4.4. Limitations and conclusions

The results of the current study should be interpreted in light of certain limitations. First, the data were derived from a relatively small sample and no prior power analyses where conducted to determine the sample size necessary to detect significant effects. Thus, it remains possible that further effects of schizotypal personality traits and self-monitoring errors on ToM performance could have emerged with a larger sample. It must be noted that to explore whether the results of the main analyses represented true effects as opposed to non-conclusive findings and to reduce the probability of type II errors, Bayesian statistics were obtained. Although Bayesian analyses provided overall support in favour of the study's findings, they also highlighted the need to test these with larger samples. For instance, although in accordance to the main analyses, results of Bayesian analyses did not find evidence for the interpersonal scale of the SPQ as a predictor of MASC hypermentalizing errors for emotions, the Bayesian evidence against including it as a predictor were also very small, thus the possibility that an association between the two could have emerged with a larger sample cannot be excluded Second, it must be noted that due to practical considerations the measures used at baseline (i.e. SPQ, self-monitoring task) differed from those used at follow-up (i.e. MASC) and this may have impacted the results of the current study. For instance, because the MASC was not

administered at baseline, it was not possible to statistically account for the effects of baseline ToM skills when assessing the effects of schizotypal traits and self-monitoring on ToM performance at follow-up. In a similar vein, although a prospective effect of self-monitoring difficulties on different forms of ToM dysfunction was identified, not having assessed ToM at baseline precluded the possibility to examine whether self-monitoring difficulties during adolescence accounted for developmental changes in ToM over time. Finally, the current study did not account for the confounding effects of interpersonal factors known to influence the development of ToM processes, such as childhood trauma (Weijers et al., 2018) and insecure attachment (Debbané, Salaminios et al., 2016).

Despite these limitations, this study is the first to prospectively examine the effects of schizotypal trait dimensions and self-monitoring difficulties during adolescence, on different types of ToM dysfunction, using an ecologically-valid experimental task. Present findings highlight the importance of encompassing evaluations of self-monitoring difficulties when clinically assessing the risk for the development of aberrant explanations of social experiences in young people presenting with schizotypal features. Within this context, mentalization- and metacognition-oriented psychotherapies (Brent & Fonagy, 2014; Debbané et al., 2016; Lysaker et al., 2020) may be applied early to support the recovery of a coherent sense of personal agency and sustain mentalizing functioning in youths who are at trait-risk for clinical psychosis.

Chapter 6

General Discussion

6.1. Summary of the main aims and findings of the current thesis

The main aims of the current thesis were to: (1) use experimental tasks and self-report measures to examine the nature of associations linking schizotypal trait dimensions to mentalizing processes involving both the understanding of the self and others during adolescence; (2) further understand the psychological processes that sustain self-awareness during adolescence by examining the effects that developmental (i.e. age), cognitive (cognitive effort) and affect-based variables (emotional valence) exert on self- and reality-monitoring performance; and (3) prospectively assess the nature of the relation between mentalizing processes sustaining self- (self-monitoring) and other-awareness (ToM) from adolescence to young adulthood.

To address these aims the empirical studies of the current thesis assessed schizotypal traits and mentalizing performance in samples of typically developing adolescents and young adults. First, the inclusion of non-clinical young people can increase our knowledge regarding factors that may impact on normative performance in psychological processes sustaining self- and other-awareness during the critical developmental period of adolescence. Furthermore, from a clinical standpoint, assessing the associations between trait-risk for psychosis and mentalizing performance in typically-developing youths can better inform the application of early prevention strategies compared to research that is limited in samples that are already at the endpoints of the psychosis-liability continuum (i.e. CHR samples).

Study 1 (Chapter 2) investigated, for the first time, the links between schizotypal trait features and self-reported mentalizing, as well as the mediational role of mentalizing on the relationship between schizotypal traits and thought problems that lie in a continuum with clinically-relevant symptoms. To this end study 1 adopted a clinically-relevant approach by (a)

investigating schizotypal traits at the individual trait-feature level (rather than measuring total dimension scores); (b) using an easy-to-administer self-report measure of mentalizing abilities that can be utilised within clinical settings for screening and routine outcome evaluation purposes (i.e. RFQ); and (3) assessing thought problems previously linked to proximal risk for psychotic disorders.

Studies 2 and 3 used novel experimental tasks to investigate young peoples' performance on two key psychological self-referential processes, namely self-monitoring (Chapter 3) and reality-monitoring (Chapter 4), as well as their linear associations with schizotypal trait dimensions. Self- and reality-monitoring were chosen as a focus of investigation because of previous research suggesting that these are consistently impaired in samples of individuals who exhibit clinical psychotic states, as well as in non-clinical adults who report psychotic-like experiences. To better understand the process of adolescent mentalizing activity as it pertains to the monitoring of self-generated activities, studies 2 and 3 also assessed self- and reality-monitoring performance between different age-groups under different conditions of cognitive effort and emotional valence respectively.

Finally, study 4 (Chapter 5) assessed ToM performance in a sample of adolescents and young adults by using an ecologically valid task, geared to assess both hyper- and hypo-mentalizing errors for complex mental state inferences involving other peoples' thoughts and feelings. The aims of the study were to prospectively assess: (a) the predictive value of schizotypal personality traits on ToM performance from adolescence to young adulthood; (b) the predictive effects of specific types of self-monitoring misattributions (externalizing vs internalizing) on different types of ToM dysfunction (hypermentalizing vs hypomentalizing); and (c) predictive interaction effects between schizotypal trait dimensions and self-monitoring difficulties on ToM performance. The later investigation was undertaken on the basis of recent conceptual models suggesting that disruptions in self-oriented reflective processing,

particularly during the premorbid stages of psychosis expression, may prospectively relate to the aberrant explanations of social experiences typically characterising core psychotic phenomena (i.e. delusions and hallucinations) (Brent et al., 2014; Debbané, Salaminios, et al., 2016).

Overall, the findings of the current thesis provide novel data suggesting that the interpersonal dimension of schizotypal personality traits (as assessed by the SPQ) which encompasses manifestations of social anxiety, constricted affect and lack of close friends, is associated with difficulties in self and other understanding during adolescence. Interestingly, contrary to previous studies in non-clinical adults (Langdon & Coltheart, 1999; Pickup, 2006; Humpston et al., 2017; Peters et al., 2007), no associations were identified in the current thesis between mentalizing difficulties and the cognitive-perceptual dimension of schizotypal traits. From a conceptual standpoint, the observed link between the interpersonal dimension of schizotypal traits and aberrant mentalizing patterns in adolescence can inform our understanding regarding early stages of a psychopathological developmental process that may contribute to increase vulnerability for psychotic disorders. These findings may also have important clinical implications for the application of early prevention psychotherapeutic treatments aiming to attenuate psychosis risk in young people who present with schizotypal trait manifestations, prior to the development of clinically-relevant symptoms.

The current thesis also offers original findings regarding the development of psychological processes sustaining self-awareness. First, in line with previous studies (Folley et al., 1983; Markham et al., 1991), the current data suggest that both self- and reality-monitoring abilities may be established in pre-adolescent development. Importantly however, results of the current thesis also suggest that reality-monitoring capacities for emotionally-charged events may undergo further elaboration from late adolescence to young adulthood. Regarding the effects of cognitive and affect-based factors on self-referential evaluations

during adolescence, in accordance with empirical studies in adult samples (Johnson et al., 1988; Suengas and Johnson, 1988; Johnson et al., 1996) results of the current thesis indicate that conditions that increase cognitive effort and emotional valence during memory encoding may lead to self- and reality-monitoring confusions respectively.

Finally, the findings of the current thesis suggest that different types of self-monitoring misattributions in adolescence may prospectively predict specific patterns of ToM dysfunction at 5-year follow-up. These findings confirm previously reported links between self- and other-awareness (Fisher et al., 2008) and highlight the importance of psychotherapeutically supporting the recovery of a coherent sense of self and personal agency, particularly among young people who present with schizotypal personality traits, to prevent the emergence of aberrant explanations of social experiences.

In the following sections of this concluding chapter, the conceptual implications of the empirical studies undertaken will be presented collectively, along with their clinical implications for the assessment and preventative treatment of young people who are at trait risk for psychosis. Furthermore, methodological issues and other limitations of the current thesis will be considered and proposed directions for future research will be presented.

6.2. Conceptual implications of the main findings

As mentioned in Chapter 1, contemporary research suggests that clinical psychosis is distally linked with schizotypal trait expression (Chapman et al., 1994; Kwapil et al., 2013), and more proximally with the breakdown of mentalizing processes sustaining the awareness of the self and others (Kim et al., 2011) (Fonagy & Target, 1996). Importantly, an increasing body of empirical research indicates that subtle mentalizing difficulties are already observable in non-clinical individuals who report high psychometric schizotypy, prior to the emergence of clinically-relevant symptoms, suggesting a pathway towards illness expression (Langdon &

Coltheart, 1999; Pickup, 2006). Although these findings are suggestive of a relationship between mentalizing dysfunction and early signs of psychosis risk, they involve adult participants either within or beyond the critical period of illness onset. This is surprising given that the first trait manifestations of psychosis risk commonly emerge during early adolescence and adopt a clinically-relevant quality during the later stages of adolescent development (Gooding et al., 2005). In addition, emerging data from neuroimaging and behavioural studies suggest that adolescence constitutes a critical period for the developmental elaboration of mentalizing processes (Choudhury et al., 2006; Dumontheil et al., 2010; Vetter et al., 2013). Most importantly perhaps, robust mentalizing abilities have been shown to support better clinical and functional outcomes in young people who report psychotic experiences (Braehler and Schwannauer, 2012; Bartels-Velthuis et al., 2011).

Overall, these different strands of research suggest that adolescence represents a critical developmental window for the preventative application of clinical interventions aiming to support mentalizing development in young people confronted with emerging risk for psychosis. However, at present, little is known about the process of adolescent mentalizing or the nature of its associations with schizotypal personality traits previously linked to increased vulnerability for the development of psychotic disorders. The current thesis used experimental tasks and self-report measures to examine mentalizing in adolescence and address the following conceptual questions: (1) What is the nature of associations linking schizotypal personality traits to mentalizing processes in adolescence? (2) What are the effects of age, cognitive effort and emotional valence on the capacity to monitor self-generated mental events in adolescence? (3) What are the prospective links between psychological processes sustaining self- (i.e. self-monitoring) and other-understanding (i.e. ToM)?

6.2.1. What is the nature of associations linking schizotypal personality traits to adolescent mentalizing?

A small body of evidence in non-clinical adult samples suggests that mentalizing difficulties involving both the understanding of the self and others relate to cognitive-perceptual trait manifestations of schizotypy (i.e. hallucination and delusion-like phenomena) (Langdon & Coltheart, 1999; Pickup, 2006; Humpston et al., 2017; Peters et al., 2007) . Studies 1-4 used different assessment methodologies to empirically examine the associations between schizotypal trait dimensions and mentalizing processes in adolescence. More specifically the studies of the current thesis assessed the associations between schizotypal trait dimensions and (a) self-reported mentalizing; (b) self- and reality-monitoring performance; and (c) ToM performance at 5-year follow-up.

Of particular interest, contrary to studies in adult populations, the empirical studies undertaken in the current thesis showed for the first time that critical associations may take place in adolescence between mentalizing abilities and the interpersonal dimension of schizotypal personality traits (i.e. trait manifestations of social anxiety, constricted affect and lack of close friends). For example, study 1 showed that schizotypal features pertaining to social anxiety, but also disorganised speech, were significantly associated with young people's self-reported mentalizing uncertainty (i.e. increased uncertainty when making mental state inferences on the basis of one's own and others' behaviours). In addition, results from study 1 showed that the effects of social anxiety and disorganised speech on thought problems relevant for psychosis may depend, at least in part, on the level of mentalizing dysfunction. Study 2 found that the interpersonal dimension of schizotypal traits in adolescence related to disruptions in self-monitoring processes involving the ability to discriminate between self-generated imagined and enacted events (i.e. discriminate in memory between what one imagined saying from what one said). Contrary to this, results of study 3 did not find significant

associations between schizotypal trait dimensions and reality-monitoring performance (i.e. discriminating between self-generated and experimenter-produced speech). Finally, study 4 showed that negative schizotypy interacted with self-monitoring difficulties in adolescence to predict ToM dysfunction in terms of hypermentalizing misattributions for cognitions at 5-year follow-up.

The association between the interpersonal dimension of schizotypy and mentalizing difficulties reported in the current thesis is a novel finding. Overall, most studies that have assessed the relationship between schizotypy and mentalizing in non-clinical samples have focused on cognitive-perceptual manifestations by only using measures of hallucination and delusion-like phenomena. Indeed, cognitive-perceptual manifestations of schizotypy are commonly used in research as outcome measures aiming to assess sub-clinical psychotic symptomatology in community samples. This however may neglect the dual nature of schizotypy, particularly as it pertains to its interpersonal dimension, both as a trait indicator of sub-clinical psychotic symptomatology, but also as an enduring personality basis that may also have a deleterious impact on young peoples' psychosocial development (see chapter 1.3.5).

Given that mentalizing processes in adolescence commonly develop against the background of close interpersonal relationships with others (Fonagy et al., 2002), one possibility is that reduced interpersonal contact in the context of negative schizotypal trait expression may contribute to undermine the elaboration of mentalizing during adolescence (Debbané & Barrantes-Vidal, 2014). For instance, young people who avoid interpersonal contact in the context of social anxiety, may deprive themselves of the opportunity to have their affective states reflected or "mirrored" by close others, thereby disrupting their capacity to construe second-order-representations of their own and other people's inner states (Ballespí et al., 2018). Similarly, restricting the range and expression of their emotional responses (i.e. constricted affect), or avoiding intimate relationships with close others (i.e. lack of close

friends) may prevent young people from effectively communicating their internal imaginative states (i.e. thoughts and feelings) within the domain of interpersonal relationships, thus disrupting their ability to "find" their psychological self in the minds of others (Brent., 2009).

Alternatively, or in conjunction, difficulties in understanding oneself and others may significantly interfere with young people's ability to function within social situations, leading them to withdraw from interpersonal contact. This is important as the widening of intimate relationships with others during adolescence typically evokes novel and distressing affective experiences (i.e. anger, despair, shame), which are often difficult to tolerate and subject to avoidance (Fonagy, 2008). Under favourable circumstances, adolescents with relatively robust mentalizing capacities manage to regulate these new experiences and integrate them in their developing self-structure. However, for some young people, disruptions in the capacity to recognise their own and other peoples' mental states can amplify the distress evoked within interpersonal situations (Debbané, Salaminios et al., 2016). For instance, difficulties in discriminating what one imagined saying or doing from what one said or did can generate experiences of increasing confusion within social situations. In a similar vein, disruptions in the capacity to make inferences about other peoples' thoughts and intentions may create difficulties in forming predictions about their behaviours, which may instead be experienced as surprising and approached with increasing suspicion (Brent, 2009). These adverse interpersonal experiences may in turn lead to the emergence of clinically relevant schizotypal outcomes such as interpersonal withdrawal, lack of intimacy, constricted affect and avolition, which signify maladaptive attempts to attenuate the distress generated within social interactions (Gumley & Schwannauer, 2006). This would resonate with data suggesting that mentalizing deficits are linked with the level of interpersonal withdrawal in patient samples (Frith, 2014; Sergi et al., 2007; Weijers et al., 2018).

Conceptually, disruptions in mentalizing processes have traditionally been linked to insecure attachment relationships with caregivers during early childhood (Fonagy et al., 2002). In a similar vein, it has been suggested that early attachment adversity, particularly as it pertains to experiences of childhood trauma and neglect, may contribute to the breakdown of mentalizing abilities in patients suffering with psychosis and CHR adults (Brent and Fonagy, 2014; Brent, 2009; Debbané, Salaminios et al., 2016). On the basis of the current findings, another possibility is that among young people who are at neurogenetic risk for psychosis, the expression of negative schizotypal traits during adolescence may also contribute to undermine the normative elaboration of mentalizing processes to increase psychosis vulnerability, even in the absence of overt attachment trauma or adversity in childhood (Debbané & Barrantes-Vidal, 2014; Debbané, Salaminios et al., 2016). Indeed, Boldrini et al. (2020a) recently reported that mentalizing difficulties, but not insecure attachment, prospectively predicted transition to clinical psychosis in a sample of CHR young adults. From a dynamic standpoint, it can be hypothesised that the process of emerging psychotic pathogenesis, as reflected by the expression of schizotypal personality traits, may actively inhibit the capacity to form emotional links with others, which in turn undermines the developmental capacity to think about, reflect on and regulate one's own inner experiences (Bion, 1959). Although an increasing number of studies suggest that the expression of negative schizotypy is prospectively linked with a heightened risk for the development of psychotic disorders (Flückiger et al., 2016; Erlenmeyer-Kimling et al., 1993), the underlying mechanisms remain incompletely understood. The results of the current thesis tentatively suggest that mentalizing dysfunction may represent a potential pathway through which negative schizotypal traits increase psychosis vulnerability during adolescence.

Surprisingly, no associations were found in the current thesis between the cognitive-perceptual dimension of schizotypy and mentalizing difficulties. This is in contrast to previous

cross-sectional data in adult samples that have shown inverse correlations between positive schizotypy and performance in both ToM and self-/reality-monitoring tasks (Pickup, 2006; Humpston et al., 2017; Peters et al., 2007). A number of methodological differences may account for the discrepancy between the current findings and those of previous studies. First, the current thesis only investigated linear associations between mentalizing performance and schizotypal trait dimensions. Conversely, the majority of previous studies have reported group-differences in mentalizing performance between people with high scores on measures of positive schizotypy or hallucination-proneness and those with low scores on these measures (Langdon and Coltheart, 1999; Larøi et al., 2004; Collignon et al., 2005; Versmissen et al., 2007). Therefore, one possibility is that the link between mentalizing difficulties and positive schizotypal manifestations may only be evident among individuals that are towards the end-points of the schizotypy continuum (i.e. in people who exhibit very high psychometric schizotypy) (Alderson-Day et al., 2019). Indeed, this may have also accounted for the unexpected lack of associations between positive schizotypy and reality-monitoring performance in study 3 (chapter 4) of the current thesis.

Another reason that could have accounted for the lack of associations between mentalizing difficulties and positive schizotypal manifestations pertains to the inclusion of adolescent samples in the current thesis. More specifically, it is possible that cognitive-perceptual manifestations and psychotic-like experiences represent outcomes, rather than predictors of mentalizing dysfunction and linear associations between the two only emerge after adolescence. In contrast, as mentioned above, schizotypal traits that specifically impact on interpersonal communication during adolescence (i.e. negative schizotypy) may also contribute to disrupt mentalizing abilities, thus their association is already evident during adolescent development. Given that the empirical studies of the current thesis are the first to examine the associations between either self- or reality-monitoring and schizotypy in typically

developing young people and no other study has examined the latter's associations with self-reported mentalizing (e.g. using the RFQ), the current findings need be interpreted with caution.

Indeed, in line with research in adult samples, the only previous study that has tested the relationship between ToM performance and schizotypy in community adolescents did find an association with the cognitive-perceptual dimension (Barragan et al., 2006). It must be noted however that similarly to studies in adult samples, Barragan et al. (2006) only reported a cross-sectional association between two. Given the lack of prospective predictive effects of positive schizotypy on ToM performance in the current thesis, it remains possible that cognitive-perceptual manifestations of schizotypy represent outcomes, rather than predictors of ToM dysfunction (Langton and Coltheart, 1999).

6.2.2. What are the effects of age, cognitive effort and emotional valance on the capacity to monitor self-generated mental events?

According to contemporary conceptualizations, psychotic experiences in clinical and non-clinical populations are associated with confusions in the identification and monitoring of self-generated information (Allen et al., 2007; Blakemore et al., 2003). Indeed, a large body of research evidence suggests that impairments in two key cognitive processes implicated in sustaining a coherent sense of self and agency, namely self-monitoring (i.e. the capacity to discriminate between what one did or said and what one imagined doing or saying) and reality-monitoring (i.e. the capacity to discriminate self-generated from other-produced events) may underpin the reality distortions characterising psychotic phenomena. Despite this evidence, the development of self-and reality-monitoring abilities during the critical period of adolescence, as well as the factors that may contribute to disrupt these remain relatively unexplored.

Interestingly, a small number of studies have shown that increased cognitive effort and emotional valence at the time of testing may disrupt self- and reality-monitoring performance and their effects are particularly pronounced among people who report psychotic experiences (Bentall et al., 1991; Larøi et al., 2004). Studies 2 and 3 of the current thesis used novel experimental tasks to empirically examine self- and reality-monitoring in adolescence by comparing performance between different age-groups, under different conditions of cognitive effort and emotional valence respectively.

First, results of the current thesis did not find a main effect of age on self- and reality-monitoring performance for speech during adolescence. This is in line with previous research in children and young adults suggesting that the capacity to monitor the source of self-generated mental events may be established prior to adolescence, during childhood development (Foley et al., 1983; Sussman, 2001). Importantly however, late adolescents displayed more reality-monitoring errors compared to young adults for words that were negatively charged. This is interesting given that no such effect was found for neutral or positively-charged material. Emerging neuroimaging data suggest that although the prefrontal cerebral networks that typically sustain basic reality-monitoring processes are already established by late childhood, their functional specialization undergoes further elaboration during adolescent development (Lagioia et al., 2011). This coincides with ongoing functional maturation in the prefrontal cortex as it pertains to its activity in the regulation of negative emotions (Young et al., 2019; Ahmed et al., 2015). Thus, it is possible that when young people undertake reality-monitoring judgments for emotionally-charged material that implicitly evoke negative arousal, the neural demands imposed on the prefrontal cortex may exceed its functional maturation, leading to disruptions in reality-monitoring performance. This effect may be particularly exacerbated during late adolescence, which entails novel emotional challenges pertaining to processes of individuation and identity formation, including the

initiation and increasing importance of romantic relationships, reduced dependence on family support and new academic pressures (Casey et al., 2010). Together, these experiences can augment young peoples' reactivity towards negative stimuli, thus increase the likelihood to confuse between self- generated and other-produced events.

Further analyses showed that the observed difference in reality-monitoring performance for negative material between late adolescents and young adults was driven by an increased rate of externalizing misattributions in the former group (i.e. confusing self-generated words as experimenter-produced). It appears that the propensity to misattribute self-generated negative material to external agents represents a normal aspect of human variation that attenuates from late adolescence towards young adulthood. This is important given that the specific externalizing pattern for negative material has also been identified among patients suffering with schizophrenia and hallucination-prone adults (Larøi et al., 2004). On the basis of the current findings, it can be tentatively hypothesised that the propensity to externalize negative material in psychosis may reflect the breakdown of neurobiological and psychosocial processes that normally sustain the developmental elaboration of reality-monitoring from late adolescence to young adulthood.

In terms of cognitive effort, a small body of research in hallucination-prone adults and young people has shown that conditions which increase the cognitive processing required during memory encoding lead to externalizing self-monitoring misattributions (i.e. confusing something one imagined saying or doing for something one did or said) (Larøi et al., 2004; Debbané et al., 2010). Contrary to these samples, the current thesis showed that typically developing adolescents display the opposite self-monitoring pattern under conditions of high cognitive effort, namely an increased rate of internalizing misattributions (i.e. confusing overt speech for silent). Conditions that increase the cognitive operations required during the encoding of overt speech typically increase its similarity to inner speech, thus leading to

confusions between the two (Johnson et al., 1993). At the same time, increased cognitive effort during the production of overt speech may reduce the encoding of its sensory-motor feedback (i.e. vocalizations, sounds), thus generating uncertainty in memory about its real and perceptually-based nature (Morosan et al., 2018). On the basis of the current data, it appears that the propensity to misattribute overt actions as imagined under conditions of increased cognitive effort constitutes a normal aspect of young people's development. Conversely, given its links with hallucination proneness, the exacerbation of externalizing self-monitoring errors under high cognitive effort may constitute an early marker of psychosis vulnerability during adolescence (Debbané et al., 2010).

6.2.3. What are the developmental links between psychological processes involved in self- and other-understanding during adolescence?

Overall, research suggests that disturbances in thinking abilities pertaining to both self- (i.e. self-monitoring) and other-understanding (i.e. ToM) are present in schizophrenia and independently contribute to the developmental unfolding of the illness during its premorbid and prodromal stages (Brent and Fonagy, 2014). Importantly however, self- and other-referential processes are not independent from each other (Frith and Frith, 1999). For instance an impaired understanding of one's inner states is likely to lead to a distorted view of other people's thoughts and intentions. Indeed, behavioural and neuroimaging studies have shown correlated task performance and shared neural activation patterns between self-monitoring and ToM evaluations among both clinical and non-clinical samples (Fisher et al., 2008 ; Murphy et al., 2010; Shad et al., 2011 ; Brunet-Gouet and Decety, 2006 ; Vinogradov et al., 2006; Simons et al., 2006; Murphy et al., 2010). These findings suggest that impairments in self-monitoring and ToM processes may reflect different dimensions of a more fundamental disturbance pertaining to the recognition of the boundaries between external and internal reality

(Murphy et al., 2010). The current thesis extends previous findings by showing that specific types of self-monitoring misattributions are prospectively linked to different forms of ToM dysfunction. More specifically, study 4 showed that externalizing and internalizing self-monitoring misattributions during adolescence respectively predicted hypermentalizing and hypomentalizing ToM for cognitions at 5-year follow-up.

First, the current findings suggest that self-monitoring confusions in adolescence may constitute early markers for the development of abnormal ToM patterns. This provides initial support to clinical conceptualizations suggesting that impaired self-awareness during the premorbid stages of psychosis expression may contribute to the elaboration of aberrant explanations of social experiences typically linked to the emergence of clinical psychotic states (i.e. paranoid delusions and hallucinations) (Brent and Fonagy, 2014). Indeed, results from the current thesis showed that when self-monitoring was impaired, the expression of negative schizotypal manifestations was prospectively linked with a propensity to overattribute thoughts and intentions in others (i.e. hypermentalizing). It is possible that in the absence of adaptive relationships with close others, difficulties in the monitoring of self-generated mental events may lead young people to adopt a hypermentalizing ToM pattern as a maladaptive means to navigate a social world of increasing complexity. Another possibility is that self-monitoring confusions may disrupt young peoples' functioning within social situations (Salaminios et al., 2020b). This may lead them to withdraw from interpersonal contact, thus preventing them from using close relationships with others to elaborate their ToM abilities (Fonagy et al., 2002). Given that negative schizotypy, externalizing self-monitoring and hypermentalizing ToM patterns have all been implicated in the expression of clinical psychosis, the current findings can inform our understanding regarding early psychopathological interactions that may contribute to increase psychosis vulnerability during adolescent development.

Furthermore, the observed associations between self-monitoring and ToM processes can extend our understanding regarding phenomenological aspects of aberrant mentalizing patterns. It has previously been suggested that individuals who experience disruptions in their mentalizing capacities, including those who experience psychotic phenomena, may turn to "pre-reflective" modes of self-organization in order to make sense of themselves in relation to others in the world (Fonagy et al., 2002 ; Debbané, Salaminios et al., 2016 ; Salaminios and Debbané, 2021). These include the *psychic equivalence* (internal reality is experienced as isomorphic to external reality) and *pretend* modes of functioning (internal reality is experienced as dissociated from external reality) (Fonagy & Target, 1996). The current data can increase our understanding regarding the underlying psychological mechanisms sustaining these pre-reflective forms of self-organization. For instance, both externalizing self-monitoring and hypermentalizing ToM patterns reflect attributional styles in which one's internal imaginative states impinge on the perception of external reality (i.e. what one thinks inside becomes real outside). In the case of the self this involves the misattribution of mental events (i.e. inner speech) as real, while in the case of others it entails the over-attribution of inner states to other peoples' overt behaviours. Therefore, the observed relation between the two may be indicative of a more general tendency to function in psychic equivalence when making attempts to understand oneself and others. Conversely, internalizing self-monitoring and hypomentalizing ToM patterns both reflect the tendency to decouple inner imaginative states from the perceptually-based domain of external reality. In terms of self-monitoring this entails the misattribution of overt actions as merely imagined, while in terms of ToM it involves the reduced attribution of inner states to other peoples' behaviours. Within this context, their association may reflect the propensity to engage in pretend mode functioning when faced with the complexity of making sense of oneself and others.

In summary, the observed associations between different types of self-monitoring and ToM patterns can shed light on the higher-order cognitive processes underpinning the phenomenology of "pre-mentalistic" forms of self-organization (Fonagy & Bateman, 2016 ; Fonagy et al., 2002). Within this context, aberrant self-monitoring and ToM patterns (e.g. externalizing self-monitoring and hypermentalizing) may not constitute static neurocognitive "deficits", as often described in the literature (Frith, 1992). Rather they may reflect compensatory, albeit maladaptive, self-regulatory attempts seeking to sustain a minimal sense of self-organization in the context of increasing complexity (Fonagy & Bateman, 2016 ; Salaminios and Debbané, 2021). Furthermore, the relation between self-monitoring and ToM in the current thesis supports the view that mentalizing represents a multifaceted neurocognitive process that is comprised of different interconnected and flexible components (e.g. self-monitoring, ToM) that may relate with each other in complex ways to determine the development of characteristic thinking patterns.

6.3. Clinical implications of the main findings

The clinical implications of the current findings lie primarily within informing early identification and prevention clinical strategies to attenuate psychosis vulnerability in young people who present with schizotypal personality traits, prior to the development of clinical symptoms. Most specifically, the data of the current thesis (1) highlight the importance of encompassing evaluations of mentalizing abilities in the assessment of psychosis risk during adolescence; and (2) provide the empirical basis supporting the preventative application of mentalization-based psychotherapy (MBT) to sustain mentalizing development in young people confronted with trait risk for psychosis.

Previous research suggests that the transition to a diagnosable form of clinical psychosis is commonly associated to the emergence of significant impairments in functional outcomes (e.g. interpersonal relating, occupational functioning), which often remain stable despite symptomatic improvement following psychotherapeutic or pharmacological treatment (Armando, Hustebaut and Debbané, 2019; Hamm et al., 2018). Similarly, young adults who display state manifestations of psychosis risk (i.e. CHR), whether they transition to a first episode of psychosis or not, do not generally display good clinical or functional outcomes (Simon et al., 2011; Carrión et al., 2013; Addington et al., 2011). Because of the above, the focus of current treatment efforts is progressively shifting towards a more preventative approach, seeking to identify and treat emerging psychosis at its earliest stages, prior to the development of clinically-relevant symptoms (Armando, Hustebaut and Debbané, 2019; McGorry et al., 2007).

Schizotypal trait manifestations commonly emerge in adolescence and reflect the earliest premorbid signs of psychosis risk. As such, the application of preventative treatments in youths who report schizotypal features may be warranted. Importantly however, the psychotic nature of schizotypal manifestations, particularly as it pertains to the interpersonal and disorganised dimensions, can often be difficult to recognise by clinicians, but also by young people and their families alike (Armando, Hustebaut and Debbané, 2019). Furthermore, questions of early treatment among young people who exhibit premorbid or prodromal manifestations can be controversial as the rates of those who do not transition to clinical psychosis are high (Addington et al., 2011; Debbané et al., 2014). Indeed, early intervention and prevention clinical practice is most often undertaken under conditions of increased uncertainty pertaining to the evolution of symptoms in those who present with sub-clinical manifestations (Armando, Hustebaut and Debbané, 2019). This highlights the need to uncover

identifiable psychological markers that can support the clinical assessment and treatment of psychosis vulnerability in young people who report schizotypal personality traits.

Mentalizing difficulties are transdiagnostic and have been implicated in the emergence of various psychiatric illnesses, such as BPD (Bateman & Fonagy, 2010); depression (Taubner et al., 2011) and functional somatic disorders (Luyten et al., 2012). Regarding the development of psychotic disorders, the current research suggests that mentalizing dysfunction represents a potential pathway through which schizotypal personality traits may increase psychosis vulnerability during adolescence. Indeed the current thesis showed that self-reported mentalizing mediated the links between schizotypal features and thought problems relevant for psychosis. Furthermore, the studies of the current thesis showed that the expression of negative schizotypal manifestations in adolescence was associated with self-monitoring confusions and the development of aberrant ToM patterns typically linked with an elevated risk for psychotic illnesses. Therefore, the current data suggest that the assessment of mentalizing abilities pertaining to the understanding of both self and others can support the evaluation of psychosis risk in young people who present with schizotypal trait manifestations. To this end, the measurement of mentalizing abilities through self-report appears to be an effective and easy-to-administer method that can be used within the context of early prevention clinical services and school settings for the purposes of screening and outcome evaluation (Fonagy et al., 2016; Salaminios et al., 2020a).

Although more research is needed to draw causal conclusions about the pathogenic impact of mentalizing dysfunction, the current findings suggest that therapies targeting mentalizing abilities may be implemented preventatively to attenuate psychosis vulnerability among young people who are faced with trait risk for the illness (Debbané, Salaminios et al., 2016; Armando, Hustebaut and Debbané, 2019). More specifically, the current data suggest that supporting young people who withdraw from adaptive interpersonal contact in the context

of negative schizotypy sustain their mentalizing abilities may protect them from the development of aberrant thinking patterns linked to psychosis. In addition, therapeutically supporting adolescents sustain their understanding of themselves and others may help them regulate the distress they experience within social situations. This can in turn reduce their tendency to avoid interpersonal contact and enable them to use positive interactions with close others to further elaborate their mentalizing capacities (Debbané, Salaminios et al., 2016; Fonagy et al., 2002). Indeed, previous research suggests that better mentalizing skills appear to support adaptation and individuation processes in adolescents recovering from a first episode of psychosis (Braehler and Schwannauer, 2012).

Within this context, targeted interventions with a mentalizing focus may be implemented on the basis of the young person's age and level of risk. First, for young adolescents who tend to socially withdraw in the context of negative schizotypy low-intensity psychosocial interventions, delivered either at the family or school levels, can be implemented to engage them into group activities that will sustain the development of normative mentalizing trajectories. Second, for high-risk adolescents who display more enduring expressions of negative schizotypy, adapted mentalization- based psychotherapies may be applied with the aim to foster better mentalizing skills to improve long-term functional outcomes and increase resilience against clinical illness (Salaminios and Debbané, 2021). This may be particularly important towards the later stages of adolescent development, which entails biological changes and new interpersonal challenges prone to evoke negative affective arousal (i.e. hormonal changes; initiation of romantic relationships), thus constituting a critical period for the disruption of mentalizing and the onset of psychotic disorders. Importantly, given that mentalizing dysfunction appears to be transdiagnostic and implicated in various psychopathological outcomes (Fonagy and Bateman, 2016), interventions with a mentalizing

focus may also sustain resilience against the emergence of non-psychotic illnesses (e.g. BPD, mood and anxiety disorders) among young people who report schizotypal manifestations.

Interestingly, the current data showed that disruptions in psychological processes sustaining self-awareness relate to trait risk for psychosis in adolescence and are predictive of the future tendency to misunderstand social situations. This suggests that helping young people who report schizotypal manifestations sustain a coherent sense of self and personal agency should constitute a primary therapeutic target to prevent the emergence of rigid or even delusional explanations of social experiences. Within this context, treatments that are primarily focused on the elimination of specific symptoms, or seek to "train" individuals into altering maladaptive thinking patterns may be insufficient to support the recovery of a coherent sense of self (Hamm et al., 2018). Rather, in order to support the development of a stable sense of self in the face of emerging psychosis, psychotherapeutic treatments may need to help individuals construct a mental representation of their inner states (Fonagy, 2000). Indeed, mentalization and metacognition-oriented approaches to psychotherapy (Brent, 2009; Hasson-Ohayon et al., 2017; Lysaker et al., 2018) may be effective in the treatment of young people who are at trait risk for psychosis, as they specifically focus on creating an intersubjective narrative construction space, which fosters in clients the capacity to think about and reflect on their emerging thoughts, feelings and experiences, as well as synthesise these to form a more coherent sense of themselves and of themselves in relation to others.

Recently, clinical adaptations of MBT for the treatment of individuals diagnosed with clinical psychosis and those who are in the prodromal risk stage of the illness have been reported in the literature (Brent, 2009; Weijers et al., 2020; Debbané et al., 2016). In addition, studies have shown the feasibility of MBT for individuals suffering with psychosis (Weijers et al., 2016) and its effectiveness in the treatment of adolescent conditions that involve comorbid psychotic symptoms (Rossouw & Fonagy, 2012). Most importantly perhaps, data from the

first randomised controlled trial suggest that although MBT may lead to robust improvements in social functioning among people diagnosed with psychotic disorders, it appears to be less effective in those suffering with chronic forms of the illness compared to patients with a recent illness-onset (Weijers et al., 2020). Given the increasing relevance of "clinical staging" (i.e. progressively tailoring the interventions offered according to progressing symptomatology and need) to support the early targeted treatment of emerging psychosis (Raballo and Larøi, 2009; McGorry et al., 2007; Nelson et al., 2016), the current data suggest that it may be fruitful to examine the effectiveness of MBT as a preventative treatment strategy for young people during the premorbid period of psychosis expression (Armando, Hustebaut and Debbané, 2019).

6.4. Methodological issues and limitations of the current thesis

The findings of the current thesis, along with their conceptual and clinical implications need to be considered in the light of certain methodological issues. Specific methodological limitations pertaining to each individual study of the thesis have been presented in the relevant chapters. The following sections will highlight more general methodological issues that cut across the studies of the current thesis. These primarily relate to the measurement of complex multidimensional constructs, such as schizotypy and mentalizing, as well as to the sampling methods used and the effects of possible confounding factors. The next sections discuss these in detail.

6.4.1. Methodological issues in the measurement of schizotypy

Although it is generally agreed that schizotypy reflects a multifaceted construct, its core dimensions, as well as the best methods to assess these remain the subject of conceptual and

empirical debate. Indeed, self-report measures developed to assess schizotypal traits often vary in terms of the manifestations they aim to capture (i.e. single manifestations vs. wider constellations), as well as in the "severity" of these (i.e. clinical vs psychometric) (Mason, 2015). As discussed in chapter 1.3.3., two approaches have been developed to measure schizotypal phenomena in the general population. Scales developed on the basis of the clinical approach either take single symptom features or diagnostic criteria, which are less frequently endorsed in the general population, as their starting point. In contrast, measures that adhere to the psychometric approach encompass items that capture "normal" or at least normally-distributed manifestations that tend to extend beyond the psychopathological domain.

The empirical studies of the current thesis used the SPQ as the main measure of schizotypal traits. The SPQ has been formulated on the basis of the DSM-III-R diagnostic criteria for schizotypal personality disorder (Mason, 2015). As such, it constitutes a "clinical" measure of schizotypal personality manifestations that may not be continuously distributed in general population. Thus, it remains possible that low mean scores on the cognitive-perceptual dimension of the SPQ may have also accounted for the lack of associations with mentalizing difficulties across the studies of the current thesis. It must be noted however that previous studies that have employed the SPQ to assess schizotypal traits in non-clinical adult and adolescent samples have reported mean, standard deviation and range values comparable to those identified in the current thesis (Fumero et al., 2018; Badoud et al., 2013; 2015b). Therefore, the scores reported in the current thesis appear to reflect the normal variation of SPQ-measured schizotypal traits in community samples. It is important to also note that the primary aim of the current thesis was to investigate how more "pathological" aspects of schizotypal trait expression may relate to mentalizing disruptions in community adolescents, thus justifying the use of the SPQ. Given that the SPQ is one of the most commonly used measures of schizotypal traits, it's use also allows for comparisons between the current findings

and those of other studies. Future studies can test the findings of the current studies by using schizotypy scales with a less clinical content than the SPQ.

Most importantly perhaps, although factor analytic studies of the most commonly used schizotypy scales have shown that they all encompass positive, negative and disorganized dimensions, these often differ to their specific content (Grant et al., 2018). For instance, the SPQ encompasses in its negative dimension trait manifestations pertaining to social anxiety, constricted affect and lack of close friends. In contrast to this, the negative dimension of other widely used schizotypy scales, such as the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE, Mason, Claridge and Jackson, 1995) and the Wisconsin Schizotypy Scales (WSS, Chapman et al., 1976) are primarily based on more subtle features of trait physical and social anhedonia. Importantly, physical and social anhedonia in community samples have been prospectively associated with an elevated risk for the development of schizophrenia-spectrum illnesses (Kwapil et al., 2013; Miettunen et al., 2011). However, the use of the SPQ in the current study precluded the investigation of associations between mentalizing and trait features of social or physical anhedonia. In a similar vein, the disorganised dimension of the SPQ primarily captures aspects of "eccentric behaviour" (i.e. odd behaviours and odd speech), while the cognitive disorganization scale of the O-LIFE is geared to measure dimensional aspects more directly related to formal thought disorder, including problems in concentration/attention and difficulties in decision-making (Grant et al., 2018). Again the latter were not assessed in the current study. In addition, high intercorrelations between the three dimensions of the SPQ were observed across the empirical studies of the current thesis, thus raising questions regarding the use of the SPQ as multidimensional measure of schizotypy (Gross et al., 2014). For instance, given that the SPQ represents a clinical measure of SPD traits, the high intercorrelations between its scales suggest that it may have tap into more global aspects cutting across its dimensions, such as perceived distress (Oezgen and Grant, 2018).

Overall, differences in the nature and content of the schizotypy scales used across different studies may impact on the variables studied, as well as on the results obtained. Indeed, at present, it appears that different measures designed to assess multidimensional aspects of schizotypy have different factor structures (Gross et al., 2014). This highlights the importance of making the theories to be tested explicit, as well as the need to be cautious when interpreting schizotypy scores in non-clinical populations (i.e. a high schizotypy score on a clinically-informed measure of schizotypy may be interpreted differently to a high schizotypy score on a measure whose content aims to capture more normally-distributed manifestations) (Mason, 2015). Furthermore, it highlights the need to further operationalise the construct of schizotypy, as well as elucidate its complex multidimensional structure. To this end, further examining the links of different schizotypal dimensions with cognitive, behavioural and neurobiological endophenotypes, as well as with environmental stress factors (e.g. trauma, attachment) may contribute to increase our understanding of this complex construct. Importantly, this can facilitate the development of new measures that will be based on the increasing empirical understanding of the construct structure of schizotypy, rather than on historical conceptual models.

6.4.2. Methodological issues in the measurement of mentalizing

As discussed in section 1.1., mentalizing represents a complex multidimensional construct that encompasses a number of processes involved in the understanding of oneself and others, such as ToM (inferring inner states from others' behaviours), empathy (relating to the emotional states of others), mindfulness (emotional self-awareness), and self-monitoring (cognitive self-awareness) (Choi-Kain & Gunderson, Fonagy & Luyten, 2009). Given its complex and multifaceted nature mentalizing is difficult to quantify and measure.

For instance, although self-report measures of mentalizing, such as the RFQ, are brief and easy-to-administer, their scores can be influenced by aspects of social desirability and most importantly by the accuracy with which individuals can assess their own mentalizing abilities (Shaw et al., 2019; Fonagy et al., 2016). In addition, the decontextualized manner in which self-report measures assess mentalizing may not take into account the implicit interactional context in which mentalizing typically unfolds, thus reducing their ecological validity (Shaw et al., 2019). To this end, future studies may incorporate experience sampling methods to assess self-reported mentalizing as it unfolds within daily life.

A second question pertains to whether the items included in operationalised self-report measures of mentalizing adequately encapsulate the complex and dimensional nature of mental state inferences. For instance, only one out of the eight items of the RFQ directly assess the understanding of others, while the rest refer to the understanding of oneself, or of oneself in relation to others, thus raising questions about the measure's face validity (Müller, Wend, et al., 2020). Thus it remains possible that RFQ scores in the current study were primarily indicative of participant's evaluations in terms of their self-oriented mentalizing abilities. It must be noted however that the RFQ has previously been shown to correlate with measures relating to other-oriented metalizing such as perspective-taking and empathy (Fonagy et al., 2016),

Another methodological issue relates to the ability of self-report measures to effectively capture different forms of mentalizing dysfunction. For example, as originally constructed, the certainty and uncertainty scales of the RFQ aim to capture maladaptive forms of hypermentalizing and hypomentalizing respectively (Fonagy et al., 2016). Importantly however, while previous studies have shown that high scores on the uncertainty scale do relate to psychopathological manifestations in both clinical and non-clinical populations, high scores on the certainty scale have been associated with better mental health outcomes (Badoud et al.,

2015; Fonagy et al., 2016; Li et al., 2020; Müller, Wend, et al., 2020). Indeed, data from a recent psychometric study using a number of psychopathological criterion variables did not find evidence that the certainty scale of the RFQ captures maladaptive forms of hypermentalizing (Müller, Wend, et al., 2020). This was taken into account in the current thesis, which only used the certainty scale of the RFQ as a measure of adaptive mentalizing functioning.

Given that self-report measures, such as the RFQ, are not geared to distinguish between self- and other-oriented mentalizing abilities (Fonagy et al., 2016), the current thesis used experimental measures to assess these. More specifically self- and reality-monitoring were used to assess the cognitive, self-oriented aspects of mentalizing, while ToM was used as a measure of other-oriented mentalizing. It must be noted however, that self/reality-monitoring and ToM are themselves complex and multifaceted processes, thus difficult to measure. For instance, while the current thesis only examined self-monitoring for speech, other studies have assessed the monitoring of physical actions. As discussed in chapter 3, self-monitoring for speech and physical actions may differ, as the latter typically involves a richer amount of sensory-motor and spatiotemporal traces encoded in memory compared to the former (Collignon et al., 2005). Indeed, previous studies that have found associations between self-monitoring and positive schizotypy, have done so by using action-monitoring paradigms (Humpston et al., 2017; Peters et al., 2007). Thus, the use of a task that assessed the monitoring of verbal and not physical actions may have accounted for the lack of associations between self-monitoring and positive schizotypy in the current thesis. Future studies can combine the assessment of self-monitoring for both verbal and physical actions to examine their links with schizotypal manifestations in young people.

Regarding reality-monitoring, the task used in the current thesis required participants to distinguish between self- and other-generated speech in memory (i.e. *offline*). Contrary to

this, a number of task-based measures involving the experimental-simulation and manipulation of participants' actions have been used to assess subjects' capacity to undertake reality-monitoring attributions *online* (i.e. in-the-moment). These have shown that individuals who report high positive schizotypy are prone to promptly disown their self-generated speech or physical movements when slight pitch or temporal modifications are experimentally introduced to the sensory feedback they receive (Versmissen et al., 2007; Hommes et al., 2011). Given that different encoding processes are involved when making offline and online reality-monitoring attributions (Dahoun et al., 2013), the findings of the current thesis may have been influenced by the use of a task that only assessed reality-monitoring evaluations in memory.

In terms of ToM, the current study employed an experimental task geared to assess the capacity to undertake complex inferences about other peoples' inner states in the context of real-life social situations (i.e. MASC). Conversely, previous studies have mainly assessed associations between schizotypy and impairments in singular components of more basic ToM abilities (i.e. facial affect recognition, false-belief, etc) (Langdon & Coltheart, 1999; Pickup, 2006). Although tasks that capture the complexity of ToM as it unfolds within day-to-day interactions may have high ecological validity, it is possible that experimental paradigms that measure relatively simple ToM inferences are more sensitive in capturing associations with cognitive-perceptual manifestations.

Finally, it must be noted that although self/reality-monitoring and ToM were selected as the focus of investigation because of evidence suggesting that they are consistently impaired in people suffering with psychosis and hallucination-prone adults (Debbané, Salaminios et al., 2016), other processes known to capture more affect-based aspects of mentalizing, such as empathy and mindfulness were not directly assessed in the current study. This is important, as disruptions in these have also been linked with psychotic phenomena in clinical and non-clinical populations (Peters et al., 2016; Montag et al., 2007). To shed more light on how

mentalizing unfolds in the context of schizotypal trait expression during adolescence, future studies should also encompass evaluations of these processes.

Overall, difficulties in the measurement of mentalizing further raise the issue of the heterogeneity of the mentalizing construct. At present, the concept of mentalizing remains too wide, without a clear operationalization and its measurement primarily relies on the assessment of its component processes (e.g ToM, self-monitoring). Importantly however, the exact nature of the relations between these processes remains incompletely understood. In this sense, more empirical work is needed to clarify the relationship between overlapping concepts, such as metacognition ToM and self-monitoring. In the absence of clear empirical evidence based on large scale comparative studies that could support the operationalization of these processes, it is important for future studies in schizotypy or schizophrenia to undertake multidimensional measurements of mentalizing. Indeed, the heterogeneity of schizotypy and schizophrenia suggest that the assessment of different mentalizing facets may be needed to identify their exact relation to these conditions.

Furthermore, it would be important for future studies to focus on the longitudinal assessment of mentalizing during critical periods for its development. This can contribute to further our understanding in terms of the factors that may contribute to disrupt or promote its developmental acquisition. To this end, it will be important to first develop new experimental tasks of mentalizing according to different developmental ages. Indeed, while traditional false-belief tasks were originally developed to assess ToM in child samples, thus may generate ceiling effects when applied to older adolescents, tasks that are geared to assess ToM within more complex social interactions (such as the MASC) may be too difficult for younger adolescents.

6.4.3. Methodological issues pertaining to sample characteristics and the effects of possible confounding factors

Certain methodological issues pertaining to the sampling methods used may have influenced the results of the current thesis. First, the data the current thesis were derived from relatively small convenience samples and further associations could have emerged by assessing larger population samples. Nevertheless, the sample sizes of the current studies are comparable to those of previous correlational studies investigating associations between schizotypal traits and mentalizing processes in community adolescents and adults (Barragan et al., 2011; Pickup., 2006). Furthermore, to account for the small sample sizes, Bayesian statistics were used to support the interpretation of the results. It must also be noted that although the current studies used samples of typically developing adolescents who at the time of testing did not suffer from any past or present psychiatric or neurogenetic disorders, data pertaining to possible psychiatric histories in the families of participants were not collected. Thus, it remains possible that some of the young people included in the studies may have had first-degree relatives who suffered with a schizophrenia-spectrum disorder. Given that unaffected first-degree relatives of individuals suffering with psychosis typically report higher scores on measures of schizotypy (Debbané et al., 2014) and display worse mentalizing skills compared to controls (Versmissen et al., 2007), this may have influenced the current data. In a similar vein, apart from age and gender, no other sociodemographic data were collected in the current thesis. This is important as different sociodemographic factors, such as ethnicity, class, and relationship status have previously been associated with schizotypal trait expression in community samples (Goulding et al., 2009; Sharpley and Peters, 1999). Finally, it must be highlighted that high scores across different schizotypy dimensions reported in the current thesis are relative to community

samples and not representative of "true schizotypes" or individuals who might be suffering from schizotypal personality disorders.

Although the current studies did take into account a number of potential covariates, such as gender, age and IQ, other confounding factors may have influenced the results obtained. Given the multiplicity of factors previously linked with schizotypal trait expression and mentalizing skills, presenting an exhaustive list of all potential confounding variables is beyond the scope of this section. Rather, certain psychosocial factors that may be particularly pertinent to the analyses conducted in the current samples are discussed. First of all, and perhaps most importantly, the studies of the current thesis did not account for the effects of attachment when investigating the associations between schizotypy and mentalizing. This is particularly important as a number of conceptual and empirical accounts have linked attachment to the development of mentalizing abilities (Fonagy et al., 2002; Fonagy et al., 1997; Sharp & Fonagy, 2008). Indeed, it is possible that the observed associations between schizotypal traits and mentalizing disruptions in the current thesis may have been influenced by the effects of insecure attachment. It must be noted however that according to recent conceptualizations, the construct of schizotypy does not constitute an independent entity, but rather represents a dynamic and complex personality system of underlying and interacting neurocognitive emotional and behavioural factors, which among others also include attachment security (or insecurity) (Fonseca-Pedrero et al., 2020). Within this context, it can be hypothesised that schizotypal personality dimensions encapsulate within them different attachment styles (i.e. preoccupied, avoidant). For instance negative schizotypy and avoidant attachment both fundamentally entail the propensity to avoid interpersonal contact and the two have consistently been shown to correlate (Berry et al., 2006; 2007; Tiliopoulos and Goodall, 2009; Debbané, Salaminios et al., 2016). Nonetheless, future studies can benefit by either investigating the mediating/moderating effects of attachment on the relationship between

schizotypy and mentalizing, or using it as a control variable in the analyses. In addition, the effects of childhood trauma and neglect were not accounted for in the current analyses. Similarly to attachment, subjective reports of childhood trauma have been shown to relate to mentalizing difficulties in individuals suffering with psychotic disorders and high-risk samples (Weijers et al., 2018; Brent and Fonagy, 2014), thus may have also impacted on the results of the current thesis. Finally, other adolescent-specific psychosocial factors that may have impacted on schizotypy scores in the current samples, such as cannabis-use (Anglin et al., 2012) and bullying-victimization (Fung and Raine, 2012) were not assessed in the current thesis.

6.5. Directions for future research

On the basis of the current findings, as well as the methodological considerations presented in the previous sections, a number of directions for future research can be proposed. Again, an exhaustive list of these is beyond the scope of the section.

A novel finding of the current thesis is that the negative schizotypal manifestations along with other schizotypal features that impede interpersonal communication with others (i.e. odd speech) relate to disruptions in mentalizing processes while these are still in the process of development during adolescence. This finding provides the empirical basis to support further conceptual and clinical research that can inform the preventative treatment of emerging psychosis in at-risk youths, prior to the development of clinical symptoms.

First, in accordance with the behavioural data of the current thesis, future studies can test whether young people who report high expressions of negative schizotypy also exhibit aberrant neurodevelopmental trajectories for key neural networks implicated in mentalizing processes.

Second, future studies can examine the possible role of embodied regulation on the associations between negative schizotypy and mentalizing. This is important as psychotic experiences, including those that specifically manifest within the interpersonal domain, have been linked with increased arousal predisposition to stress (Clamor et al., 2015) and reduced interoceptive awareness (Koreki et al., 2021). Furthermore, it has been proposed that in the context of affective arousal during adolescence, failures to form mental representations of embodied signals to regulate their self-experience can lead young people develop aberrant explanations of social experiences as a maladaptive means to attenuate experiences of increasing confusion (Salaminios and Debbané, 2021). On the basis of the current findings, it may be fruitful to empirically examine whether difficulties in the recognition and regulation of embodied interoceptive signals among young people who report negative schizotypy may trigger the use of maladaptive mentalizing patterns.

In addition, it will be important to specify whether the interpersonal trait manifestations shown to relate to mentalizing difficulties in adolescence are underpinned by experiences of perceived threat in the context of social situations, by a diminished anticipation of reward from interpersonal relations, or by a more general lack of motivation to engage in these. This could in turn help us identify creative ways to therapeutically support young people confronted with trait risk for psychosis engage in social activities that can sustain the normative development of mentalizing and increase resilience against transition to clinical illness. In terms of the methods used to assess schizotypy, future studies should also test the current findings by employing scales geared to provide more refined assessments of negative schizotypy, including expressions of physical and social anhedonia, or scales with more "continuously-distributed" and less clinical item content than the SPQ.

From a clinical point of view, the observed association between negative schizotypal manifestations and mentalizing processes also supports the development of small-scale pilot

studies that will test the feasibility and effectiveness of preventative interventions aiming to sustain mentalizing development in the face of trait risk for psychosis during adolescence. First, studies can test whether low intensity psychosocial interventions, delivered either at the family or school levels, can successfully engage young people who exhibit negative schizotypal manifestations into social activities that will foster the development of mentalizing. Second, for high-risk adolescents who report more enduring expressions of negative schizotypy, studies can also test whether adapted mentalization- and metacognition-based psychotherapies are effective in fostering mentalizing skills to improve long-term clinical and functional outcomes. Furthermore mentalizing can be used as a process-outcome variable in studies of other psychosocial interventions applied to the treatment of psychosis risk in adolescence.

Given the observed associations between specific types of self-monitoring (i.e. externalizing vs internalizing) and ToM patterns (i.e. hypermentalizing and hypomentalizing), as well as proposed hypotheses about their potential "defensive" function, another interesting direction for future research would be to identify the conditions (i.e. interpersonal, affective) that may trigger the use of these. Indeed, it has been argued that individuals who exhibit psychotic phenomena may concurrently display both hypo- and hyper-mentalizing ToM patterns (Langdon and Brock, 2008). Similarly, young people who are at neurogenetic risk for psychosis have been found to exhibit greater rates of both externalizing and internalizing self-monitoring misattributions compared to controls (Debbané et al., 2010). Thus, identifying affective and interpersonal stressors that may, at different times, potentiate the use of one aberrant attributional pattern over the other, among the same individuals, may be clinically-relevant for the application of targeted and personalized preventative interventions during adolescence. To this end, future studies can utilise innovative methodological advances (i.e. experience sampling methods) to assess mentalizing process as they unfold in young peoples' daily lives, which will invariably entail interpersonal and affective stress.

Given the links between negative schizotypal manifestations, self-monitoring confusions and hypermentalizing ToM patterns observed in the current thesis, future research can also explore whether these interact to prospectively predict the worsening of schizotypal manifestations in community samples, or the transition to clinical psychosis in CHR samples. Furthermore, because the current data showed that self-reported mentalizing difficulties during adolescence mediated the effects of schizotypal traits on thought problems relevant for psychosis, future studies can more directly test this in CHR samples by using attenuated and intermittent psychotic symptoms as outcome variables.

As mentioned in the previous section, future studies that test the associations between schizotypal traits and mentalizing processes can also benefit by encompassing evaluations of other psychosocial risk factors known to influence these, including among others, attachment security, childhood trauma and neglect, cannabis use and bullying-victimization.

Overall, mentalizing is a useful construct for the understanding of schizophrenia and schizotypy, as empirical studies have shown its utility in terms of capturing variance in symptomatic and functional outcomes among both clinical and non-clinical samples. Importantly, a number of questions require further empirical investigation. First, the directionality of the observed associations between schizotypy and mentalizing remains unclear. Surprisingly, there is a lack of longitudinal investigations in adult and adolescent samples that would empirically identify whether mentalizing difficulties represent the outcomes of schizotypy or whether they contribute to increase schizotypy. Furthermore, the identification and differentiation of the specific mentalizing impairments that characterise the positive, negative and disorganised dimensions of schizotypy (and schizophrenia) is needed with the aim to develop personalised interventions that will attenuate the risk for the development of clinical psychosis and support better functional outcomes.

6.6. Concluding remarks

Overall, the findings of the current thesis suggest, for the first time, that the expression of negative schizotypal traits, which includes manifestations of social anxiety, constricted affect and lack of close friends, is associated with difficulties in self and other understanding during adolescence. Furthermore, self-reported mentalizing uncertainty partly mediated the effects of schizotypal manifestations pertaining to social anxiety and odd speech on thought problems relevant for psychosis. Contrary to previous studies in non-clinical adults, no associations were identified in the current thesis between mentalizing difficulties and cognitive-perceptual manifestations of schizotypy. Regarding the development of self-oriented processes involved in monitoring self-generated contents, the current data suggest that both self- and reality-monitoring abilities may be established in pre-adolescent development. Importantly however, results of the current thesis also suggest that reality-monitoring capacities for emotionally-charged material may undergo further elaboration from late adolescence to young adulthood. Regarding the effects of cognitive and affect-based factors on self-referential evaluations during adolescence, results of the current thesis indicate that conditions that increase cognitive effort and emotional valence during memory encoding may lead to self- and reality-monitoring confusions respectively. In addition, the findings of the current thesis suggest that externalizing and internalising self-monitoring misattributions prospectively predicted hypermentalizing and hypomentalizing ToM for cognitions respectively. Negative schizotypal manifestations prospectively predicted hypermentalizing ToM for cognitions at high but not low or moderate rates of externalizing self-monitoring misattributions.

The current thesis has raised new questions and pointed to new directions for further research in the study of the associations between mentalizing and schizotypal trait expression. Furthermore, the current findings have the potential to inform early identification and

prevention treatment strategies for young people who present with trait vulnerability for psychosis.

References

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*: Department of Psychiatry, University of Vermont.

Acosta, H., Straube, B., & Kircher, T. (2019). Schizotypy and mentalizing: An fMRI study. *Neuropsychologia*, *124*, 299-310.

Addington, J., Cornblatt, B. A., Cadenhead, K. S., Cannon, T. D., McGlashan, T. H., Perkins, D. O., & Heinssen, R. (2011). At clinical high risk for psychosis: outcome for nonconverters. *American Journal of Psychiatry*, *168*(8), 800-805.

Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental cognitive neuroscience*, *15*, 11-25.

Alderson-Day, B., Smailes, D., Moffatt, J., Mitrenga, K., Moseley, P., & Fernyhough, C. (2019). Intentional inhibition but not source memory is related to hallucination-proneness and intrusive thoughts in a university sample. *cortex*, *113*, 267-278.

Allen, P., Aleman, A., & Mcguire, P.K. (2007). Inner speech models of auditory verbal hallucinations: evidence from behavioural and neuroimaging studies. *International Review of Psychiatry, 19*(4), 407-415.

Allen, P., Freeman, D., Johns, L., & McGuire, P. (2006). Misattribution of self-generated speech in relation to hallucinatory proneness and delusional ideation in healthy volunteers. *Schizophrenia Research*, *84*(2-3), 281-288.

Anglin, D. M., Corcoran, C. M., Brown, A. S., Chen, H., Lighty, Q., Brook, J. S., & Cohen, P. R. (2012). Early cannabis use and schizotypal personality disorder symptoms from adolescence to middle adulthood. *Schizophrenia research*, *137*(1-3), 45-49.

Ainsworth, M. D., Blehar, M., Waters, E., & Wall, S. (1978). Patterns of attachment.

Armando, M., Hutsebaut, J., & Debbané, M. (2019). A mentalization-informed staging approach to clinical high risk for psychosis. *Frontiers in psychiatry*, *10*, 385.

Badoud, D., Speranza, M., & Debbané, M. (2016). Vers un modèle du développement des dimensions de la mentalisation a l'adolescence. *Revue québécoise de psychologie*, *37*(3), 49-68.

Badoud, D., Luyten, P., Fonseca-Pedrero, E., Eliez, S., Fonagy, P., & Debbané, M. (2015a). The French version of the Reflective Functioning Questionnaire: validity data for adolescents and adults and its association with non-suicidal self-injury. *PloS one, 10*(12), e0145892.

Badoud, D., Billieux, J., Eliez, S., Imhof, A., Heller, P., Eytan, A., & Debbané, M. (2015b). Covariance and specificity in adolescent schizotypal and borderline trait expression. *Early Intervention in Psychiatry*, *9*(5), 378-387.

Badoud, D., Billieux, J., Van der Linden, M., Eliez, S., & Debbané, M. (2013). Encoding style and its relationships with schizotypal traits and impulsivity during adolescence. *Psychiatry research*, *210*(3), 1020-1025.

Badoud, D., Chanal, J., der Linden Van, M., Eliez, S., & Debbané, M. (2011). Validation study of the French schizotypal personality questionnaire in a sample of adolescents: a confirmatory factor analysis. *L'encéphale, 37*(4), 299-307.

Ballespí, S., Vives, J., Sharp, C., Tobar, A., & Barrantes-Vidal, N. (2019). Hypermentalizing in Social Anxiety: Evidence for a Context-Dependent Relationship. *Frontiers in psychology*, *10*, 1501.

Ballespí, S., Pérez-Domingo, A., Vives, J., Sharp, C., & Barrantes-Vidal, N. (2018). Childhood behavioral inhibition is associated with impaired mentalizing in adolescence. *PloS one, 13*(3), e0195303.

Banerjee, R., & Henderson, L. (2001). Social-cognitive factors in childhood social anxiety: a preliminary investigation. *Social Development, 10*(4), 558-572.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*(1), 37-46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry*, *42*(2), 241-251.

Barragan, M., Laurens, K. R., Navarro, J. B., & Obiols, J. E. (2011). 'Theory of Mind', psychotic-like experiences and psychometric schizotypy in adolescents from the general population. *Psychiatry research, 186*(2), 225-231.

Barrantes-Vidal, N., Grant, P., & Kwapil, T. R. (2015). The role of schizotypy in the study of the etiology of schizophrenia spectrum disorders. *Schizophrenia bulletin, 41*(suppl_2), S408-S416.

Barrantes-Vidal, N., Fañanás, L., Rosa, A., Caparrós, B., Riba, M. D., & Obiols, J. E. (2003). Neurocognitive, behavioural and neurodevelopmental correlates of schizotypy clusters in adolescents from the general population. *Schizophrenia Research*, *61*(2-3), 293-302.

Bartels-Velthuis, A. A., Wigman, J. T. W., Jenner, J. A., Bruggeman, R., & Van Os, J. (2016). Course of auditory vocal hallucinations in childhood: 11-year follow-up study. *Acta Psychiatrica Scandinavica*, *134*(1), 6-15.

Bartels-Velthuis, A. A., Blijd-Hoogewys, E. M., & Van Os, J. (2011). Better theory-of-mind skills in children hearing voices mitigate the risk of secondary delusion formation. *Acta Psychiatrica Scandinavica, 124*(3), 193-197.

Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World psychiatry, 9*(1), 11-15.

Bateman, A. W., & Fonagy, P. (2004). Mentalization-based treatment of BPD. *Journal of personality disorders, 18*(1), 36-51.

Bentall, R.P., Baker, G.A., & Havers, S. (1991). Reality monitoring and psychotic hallucinations. *British Journal of Clinical Psychology, 30*(3), 213-222.

Berry, K., Band, R., Corcoran, R., Barrowclough, C., & Wearden, A. (2007). Attachment styles, earlier interpersonal relationships and schizotypy in a non-clinical sample. *Psychology and Psychotherapy: Theory, Research and Practice*, *80*(4), 563-576.

Berry, K., Wearden, A., Barrowclough, C., & Liversidge, T. (2006). Attachment styles, interpersonal relationships and psychotic phenomena in a non-clinical student sample. *Personality and Individual Differences*, *41*(4), 707-718.

Billieux, J., Rochat, L., Ceschi, G., Carré, A., Offerlin-Meyer, I., Defeldre, A.-C., Khazaal, Y., Besche-Richard, C., & Van der Linden, M. (2012). Validation of a short French version of the UPPS-P Impulsive Behavior Scale. *Comprehensive Psychiatry*, *53*(5), 609-615.

Bion, W. R. (1962). Learning from experience *London: Karnac*.

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267-277.

Blakemore, S.-J., Oakley, D.A., & Frith, C. (2003). Delusions of alien control in the normal brain. *Neuropsychologia, 41*(8), 1058-1067.

Blakemore, S. J., Smith, J., Steel, R., Johnstone, E. C., & Frith, C. D. (2000). The perception of self‑produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychological medicine*, *30*(5), 1131-1139.

Blijd-Hoogewys, E. M. A., Van Geert, P. L. C., Serra, M., & Minderaa, R. B. (2008). Measuring theory of mind in children. Psychometric properties of the ToM storybooks. *Journal of autism and Developmental Disorders*, *38*(10), 1907-1930.

Bliksted, V., Frith, C., Videbech, P., Fagerlund, B., Emborg, C., Simonsen, A., ... & Campbell-Meiklejohn, D. (2019). Hyper-and hypomentalizing in patients with first-episode schizophrenia: fMRI and behavioral studies. *Schizophrenia bulletin*, *45*(2), 377-385.

Bogren, M., Mattisson, C., Tambs, K., Horstmann, V., Munk-Jørgensen, P., & Nettelbladt, P. (2010). Predictors of psychosis: a 50-year follow-up of the Lundby population. *European archives of psychiatry and clinical neuroscience*, *260*(2), 113-125.

Boldrini, T., Pontillo, M., Tanzilli, A., Giovanardi, G., Di Cicilia, G., Salcuni, S., ... & Lingiardi, V. (2020a). An attachment perspective on the risk for psychosis: clinical correlates and the predictive value of attachment patterns and mentalization. *Schizophrenia Research*, *222*, 209-217.

Boldrini, T., Tanzilli, A., Di Cicilia, G., Gualco, I., Lingiardi, V., Salcuni, S., ... & Pontillo, M. (2020b). Personality Traits and Disorders in Adolescents at Clinical High Risk for Psychosis: Toward a Clinically Meaningful Diagnosis. *Frontiers in Psychiatry*, *11*, 1258.

Bora, E., Sehitoglu, G., Aslier, M., Atabay, I., & Veznedaroglu, B. (2007). Theory of mind and unawareness of illness in schizophrenia. *European archives of psychiatry and clinical neuroscience*, *257*(2), 104-111.

Bora, E., & Pantelis, C. (2013). Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia: systematic review and meta-analysis. *Schizophrenia Research, 144*(1), 31-36.

Braff, D. L., Geyer, M. A., Light, G. A., Sprock, J., Perry, W., Cadenhead, K. S., & Swerdlow, N. R. (2001). Impact of prepulse characteristics on the detection of sensorimotor gating deficits in schizophrenia. *Schizophrenia Research, 49*(1-2), 171-178.

Braehler, C., & Schwannauer, M. (2012). Recovering an emerging self: Exploring reflective function in recovery from adolescent-onset psychosis. *Psychology and Psychotherapy: theory, research and practice, 85*(1), 48-67.

Brébion, G., Amador, X., David, A., Malaspina, D., Sharif, Z., & Gorman, J.M. (2000). Positive symptomatology and source-monitoring failure in schizophrenia—an analysis of symptom-specific effects. *Psychiatry research, 95*(2), 119-131.

Brent, B. K. (2015). A mentalization-based approach to the development of the therapeutic alliance in the treatment of schizophrenia. *Journal of clinical psychology*, *71*(2), 146-156.

Brent, B. (2009). Mentalization-based psychodynamic psychotherapy for psychosis. *Journal of clinical psychology, 65*(8), 803-814.

Brent, B. K., & Fonagy, P. (2014). A mentalization-based treatment approach to disturbances of social understanding in schizophrenia. *Social Cognition and Metacognition in Schizophrenia: Psychopathology and Treatment Approaches eds. PH Lysaker, G. Dimaggio, M. Brune (Gurgaon: Elsevier Science and Technology)*, 245-259.

Brink, V., van Driel, C., El Bouhaddani, S., Wardenaar, K. J., van Domburgh, L., Schaefer, B., ... & Veling, W. (2020). Spontaneous discontinuation of distressing auditory verbal hallucinations in a school-based sample of adolescents: a longitudinal study. *European child & adolescent psychiatry, 29*(6), 777-790.

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: a meta-analytic review. *Psychological medicine, 43*(12), 2465-2475.

Brüne, M. (2005). "Theory of mind" in schizophrenia: a review of the literature. *Schizophrenia bulletin, 31*(1), 21-42.

Brunet-Gouet, E., Decety, J. Social brain dysfunctions in schizophrenia: a review of neuroimaging studies Psychiatry Res., 148 (2–3) (2006), pp. 75-92

Caputo, G. B. (2010). Strange-face-in-the-mirror illusion. *Perception, 39*(7), 1007

Carrión, R. E., McLaughlin, D., Goldberg, T. E., Auther, A. M., Olsen, R. H., Olvet, D. M., ... & Cornblatt, B. A. (2013). Prediction of functional outcome in individuals at clinical high risk for psychosis. *JAMA psychiatry, 70*(11), 1133-1142.

Casey, B. J., Jones, R. M., Levita, L., Libby, V., Pattwell, S. S., Ruberry, E. J., ... & Somerville, L. H. (2010). The storm and stress of adolescence: insights from human

imaging and mouse genetics. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *52*(3), 225-235.

Chapman, L. J., Chapman, J. P., Kwapil, T. R., Eckblad, M., & Zinser, M. C. (1994). Putatively psychosis-prone subjects 10 years later. *Journal of abnormal psychology, 103*(2), 171.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-image aberration in Schizophrenia. *Journal of abnormal psychology*, *87*(4), 399.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of abnormal psychology*, *85*(4), 374.

Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: Ontogeny, assessment, and application in the treatment of borderline personality disorder. *American Journal of Psychiatry, 165*(9), 1127-1135.

Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social cognitive and affective neuroscience, 1*(3), 165-174.

Chung, Y. S., Kang, D. H., Shin, N. Y., Yoo, S. Y., & Kwon, J. S. (2008). Deficit of theory of mind in individuals at ultra-high-risk for schizophrenia. *Schizophrenia research*, *99*(1-3), 111-118.

Ciaramidaro, A., Bölte, S., Schlitt, S., Hainz, D., Poustka, F., Weber, B.,Walter, H. (2015). Schizophrenia and autism as contrasting minds: neural evidence for the hypo-hyper-intentionality hypothesis. Schizophrenia Bulletin, 41(1), 171–179

Cicero, D. C., Kerns, J. G., & McCarthy, D. M. (2010). The Aberrant Salience Inventory: A new measure of psychosis proneness. *Psychological Assessment*, *22*(3), 688.

Clamor, A., Warmuth, A. M., & Lincoln, T. M. (2015). Arousal predisposition as a
vulnerability indicator for psychosis: a general population online stress induction
study. *Schizophrenia research and treatment*, *2015*.

Claridge, G. E. (1997). *Schizotypy: Implications for illness and health*. Oxford University
Press.

Claridge, G., & Beech, T. (1995). Fully and quasi-dimensional constructions of schizotypy.
In A. Raine, T. Lencz, & S. A. Mednick (Eds.), *Schizotypal personality* (p. 192–216).
Cambridge University Press

Claridge, G., & Broks, P. (1984). Schizotypy and hemisphere function—I: Theoretical
considerations and the measurement of schizotypy. *Personality and Individual
Differences*, *5*(6), 633-648.

Clemmensen, L., Van Os, J., Drukker, M., Munkholm, A., Rimvall, M. K., Vaever, M., ... &
Jeppesen, P. (2016). Psychotic experiences and hyper-theory-of-mind in
preadolescence—a birth cohort study. *Psychol Med*, *46*(1), 87-101.

Clemmensen, L., van Os, J., Skovgaard, A. M., Væver, M., Blijd-Hoogewys, E. M., Bartels-
Velthuis, A. A., & Jeppesen, P. (2014). Hyper-theory-of-mind in children with
psychotic experiences. *Plos one*, *9*(11).

Collignon, O., Van der Linden, M., & Larøi, F. (2005). Source monitoring for actions in
hallucination proneness. *Cognitive Neuropsychiatry*, *10*(2), 105-123.

Corcoran, R., & Frith, C. (2003). Autobiographical memory and theory of mind: evidence of
a relationship in schizophrenia. *Psychological medicine*, *33*(5), 897.

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophrenia research*, *17*(1), 5-13.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental science, 1*(2), 255-259.

Dahoun, T., Eliez, S., Chen, F., Badoud, D., Schneider, M., Larøi, F., & Debbane, M. (2013). Action simulation in hallucination-prone adolescents. *Frontiers in human neuroscience*, *7*, 329.

Daniel, C., & Mason, O. J. (2015). 9 Inducing psychotic-like experiences. *Schizotypy: New dimensions*, 147.

Davidson, C. A., Piskulic, D., Addington, J., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., McGlashan, T. H., Perkins, D. O., Seidman, L. J., & Tsuang, M. T. (2018). Age-related trajectories of social cognition in youth at clinical high risk for psychosis: An exploratory study. *Schizophrenia Research*.

Debbané, M., & Barrantes-Vidal, N. (2014). Schizotypy from a developmental perspective. *Schizophrenia bulletin, 41*(suppl_2), S386-S395.

Debbané, M., Benmiloud, J., Salaminios, G., Solida-Tozzi, A., Armando, M., Fonagy, P., & Bateman, A. (2016). Mentalization-based treatment in clinical high-risk for psychosis: a rationale and clinical illustration. *Journal of Contemporary Psychotherapy, 46*(4), 217-225.

Debbané, M., Salaminios, G., Luyten, P., Badoud, D., Armando, M., Tozzi, A. S., Fonagy, P., & Brent, B. K. (2016). Attachment, neurobiology, and mentalizing along the psychosis continuum. *Frontiers in human neuroscience,* 10, 406

Debbané, M., Eliez, S., Badoud, D., Conus, P., Flückiger, R., & Schultze-Lutter, F. (2014). Developing psychosis and its risk states through the lens of schizotypy. *Schizophrenia bulletin, 41*(suppl_2), S396-S407.

Debbané, M., Badoud, D., Balanzin, D., & Eliez, S. (2013). Broadly defined risk mental states during adolescence: disorganization mediates positive schizotypal expression. *Schizophrenia Research, 147*(1), 153-156.

Debbané, M., Van der Linden, M., Glaser, B., & Eliez, S. (2010). Monitoring of self-generated speech in adolescents with 22q11. 2 deletion syndrome. *British Journal of Clinical Psychology, 49*(3), 373-386.

Debbané, M., Van der Linden, M., Gex-Fabry, M., & Eliez, S. (2009). Cognitive and emotional associations to positive schizotypy during adolescence. *Journal of Child Psychology and Psychiatry, 50*(3), 326-334.

Debbané, M., Van der Linden, M., Glaser, B., & Eliez, S. (2008). Source monitoring for actions in adolescents with 22q11. 2 deletion syndrome (22q11DS). *Psychological medicine, 38*(6), 811-820.

Dewhurst, S.A. (1999). Cognitive effort and recollective experience in recognition memory. *Memory, 7*(2), 129-146.

Docherty, N.M. (2012). Missing referents, psychotic symptoms, and discriminating the internal from the externalized. *Journal of abnormal psychology, 121*(2), 416.

Docherty, N. M., & Gottesman, I. I. (2000). A twin study of communication disturbances in schizophrenia. *The Journal of nervous and mental disease, 188*(7), 395-401.

Dominguez, M. D. G., Saka, M. C., Lieb, R., Wittchen, H. U., & van Os, J. (2010). Early expression of negative/disorganized symptoms predicting psychotic experiences and

subsequent clinical psychosis: a 10-year study. American Journal of Psychiatry, 167(9), 1075–1082.

Dimaggio, G., Vanheule, S., Lysaker, P. H., Carcione, A., & Nicolò, G. (2009). Impaired self-reflection in psychiatric disorders among adults: a proposal for the existence of a network of semi independent functions. *Consciousness and cognition*, *18*(3), 653-664.

Dumas, P., Bouafia, S., Gutknecht, C., Saoud, M., Dalery, J., & d'Amato, T. (2000). Validation of the French version of the Raine Schizotypal Personality Disorder Questionnaire. *L'encéphale, 26*(5), 23-29.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental science, 13*(2), 331-338.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... & Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of autism and developmental disorders*, *36*(5), 623-636.

Eaton, W. W., Romanoski, A., Anthony, J. C., & Nestadt, G. (1991). Screening for psychosis in the general population with a self-report interview. *Journal of nervous and mental Disease*.

Eckblad, M., & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of consulting and clinical psychology*, *51*(2), 215.

Eckblad, M. L., Chapman, L. J., Chapman, J. P., & Mishlove, M. (1982). The revised social anhedonia scale. *Unpublished test*.

Erlenmeyer-Kimling, L., Cornblatt, B. A., Rock, D., Roberts, S., Bell, M., & West, A. (1993). The New York high-risk project: anhedonia, attentional deviance, and psychopathology. *Schizophrenia Bulletin*, *19*(1), 141-153.

Escher, S., Romme, M., Buiks, A., Delespaul, P., & Van Os, J. I. M. (2002). Independent course of childhood auditory hallucinations: a sequential 3-year follow-up study. *The British Journal of Psychiatry*, *181*(S43), s10-s18.

Ettinger, U., Mohr, C., Gooding, D. C., Cohen, A. S., Rapp, A., Haenschel, C., & Park, S. (2015). Cognition and brain function in schizotypy: a selective review. *Schizophrenia bulletin*, *41*(suppl_2), S417-S426.

Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and individual differences*, *6*(1), 21-29.

Ferchiou, A., Schürhoff, F., Bulzacka, E., Mahbouli, M., Leboyer, M., & Szöke, A. (2010). Source monitoring: general presentation and review of literature in schizophrenia. *L'encéphale*, *36*(4), 326-333.

Fett, A. K. J., Viechtbauer, W., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 573-588.

Finke, R.A., Johnson, M.K., Shyi, G.C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition*, *16*(2), 133-137.

Fisher, M., McCoy, K., Poole, J. H., & Vinogradov, S. (2008). Self and other in schizophrenia: a cognitive neuroscience perspective. *American Journal of Psychiatry*, *165*(11), 1465-1472.

Flückiger, R., Ruhrmann, S., Debbané, M., Michel, C., Hubl, D., Schimmelmann, B. G., ... & Schultze-Lutter, F. (2016). Psychosis-predictive value of self-reported schizotypy in a clinical high-risk sample. *Journal of abnormal psychology*, *125*(7), 923.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child development*, 1145-1155.

Foley, M.A., Johnson, M.K., & Raye, C.L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child development*, 51-60.

Fonagy, P. (2000). Attachment and borderline personality disorder. *Journal of the american psychoanalytic association, 48*(4), 1129-1146.

Fonagy, P., & Bateman, A. W. (2016). Adversity, attachment, and mentalizing. *Comprehensive psychiatry, 64*, 59-66.

Fonagy, P., Gergely, G., Jurist, E.L., & Target, M. (2002). *Affect regulation, mentalization and the development of the self*: Routledge.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and psychopathology, 21*(4), 1355-1381.

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y.-W., Warren, F., Howard, S., Ghinai, R., Fearon, P., & Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The Reflective Functioning Questionnaire. *PloS one, 11*(7), e0158678.

Fonagy, P., Redfern, S., & Charman, T. (1997). The relationship between belief-desire reasoning and a projective measure of attachment security (SAT). *British Journal of Developmental Psychology, 15*(1), 51-61.

Fonagy, P., & Target, M. (1996). Playing with reality I. *The International journal of psycho-analysis, 77*(2), 217.

Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and psychopathology, 9*(4), 679-700.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). Reflective-functioning manual, version 5.0, for application to adult attachment interviews. *London: University College London*, 161-2.

Fonseca-Pedrero, E., Ortuño-Sierra, J., Inchausti, F., Rodríguez-Testal, J. F., & Debbané, M. (2020). Beyond clinical high-risk state for psychosis: the network structure of multidimensional psychosis liability in adolescents. *Frontiers in psychiatry, 10*, 967.

Fonseca-Pedrero, E., Debbané, M., Ortuño-Sierra, J., Chan, R., Cicero, D., Zhang, L., Brenner, C., Barkus, E., Linscott, R., & Kwapil, T. (2018). The structure of schizotypal personality traits: a cross-national study. *Psychological medicine, 48*(3), 451-462.

Fotopoulou, A., & Tsakiris, M. (2017). Mentalizing homeostasis: The social origins of interoceptive inference. *Neuropsychoanalysis, 19*(1), 3-28.

Franck, N., Rouby, P., Daprati, E., Daléry, J., Marie-Cardine, M., & Georgieff, N. (2000). Confusion between silent and overt reading in schizophrenia. *Schizophrenia Research, 41*(2), 357-364.

Freeman, D. (2007). Suspicious minds: the psychology of persecutory delusions. *Clinical psychology review, 27*(4), 425-457.

Frith, C. D. (2014). *The cognitive neuropsychology of schizophrenia*: Psychology press.

Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological medicine, 34*(3), 385-389.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Psychology press.

Frith, C.D., Blakemore, S.-J., & Wolpert, D.M. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Research Reviews, 31*(2-3), 357-363.

Frith, C. D., & Corcoran, R. (1996). Exploring theory of mind in people with schizophrenia. *Psychological medicine*, *26*(3), 521-530.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*(4), 531-534.

Fumero, A., Marrero, R. J., & Fonseca-Pedrero, E. (2018). Well-being in schizotypy: the effect of subclinical psychotic experiences. *Psicothema*, *30*(2), 177-82.

Fung, A. L. C., & Raine, A. (2012). Peer victimization as a risk factor for schizotypal personality in childhood and adolescence. *Journal of personality disorders*, *26*(3), 428-434.

Fusar-Poli, P., Borgwardt, S., Bechdolf, A., Addington, J., Riecher-Rössler, A., Schultze-Lutter, F., ... & Yung, A. (2013). The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA psychiatry*, *70*(1), 107-120.

Fusar-Poli, P., de Pablo, G. S., Correll, C. U., Meyer-Lindenberg, A., Millan, M. J., Borgwardt, S., ... & Arango, C. (2020). Prevention of psychosis: advances in detection, prognosis, and intervention. *Jama Psychiatry*, *77*(7), 755-765.

Fyfe, S., Williams, C., Mason, O. J., & Pickup, G. J. (2008). Apophenia, theory of mind and schizotypy: perceiving meaning and intentionality in randomness. *Cortex*, *44*(10), 1316-1325.

Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual differences*, *30*(8), 1311-1327.

Gawęda, Ł., Li, E., Lavoie, S., Whitford, T. J., Moritz, S., & Nelson, B. (2018). Impaired action self-monitoring and cognitive confidence among ultra-high risk for psychosis and first-episode psychosis patients. *European Psychiatry*, *47*, 67-75.

Gawęda, Ł., Moritz, S., & Kokoszka, A. (2012). Impaired discrimination between imagined and performed actions in schizophrenia. *Psychiatry research, 195*(1-2), 1-8.

Gergely, G. (2004). The role of contingency detection in early affect–regulative interactions and in the development of different types of infant attachment. *Social development, 13*(3), 468-478

Gergely, G., & Watson, J. S. (1999). Early socio-emotional development: Contingency perception and the social-biofeedback model. *Early social cognition: Understanding others in the first months of life, 60,* 101-136.

Gooding, D.C., Tallent, K.A., & Matts, C.W. (2005). Clinical status of at-risk individuals 5 years later: further validation of the psychometric high-risk strategy. *Journal of abnormal psychology, 114*(1), 170.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160*(4), 636-645.

Goulding, S. M., McClure-Tone, E., & Compton, M. T. (2009). Associations between multiple dimensions of schizotypy and sociodemographic variables in a nonpsychiatric sample of young adults. *The Journal of nervous and mental disease, 197*(10), 786-789.

Grant, P., Green, M. J., & Mason, O. J. (2018). Models of schizotypy: the importance of conceptual clarity. *Schizophrenia bulletin, 44*(suppl_2), S556-S563.

Greenwood, T. A., Swerdlow, N. R., Gur, R. E., Cadenhead, K. S., Calkins, M. E., Dobie, D. J., Freedman, R., Green, M. F., Gur, R. C., & Lazzeroni, L. C. (2013). Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. *American Journal of Psychiatry, 170*(5), 521-532.

Gross, G. M., Mellin, J., Silvia, P. J., Barrantes-Vidal, N., & Kwapil, T. R. (2014). Comparing the factor structure of the Wisconsin Schizotypy Scales and the

Schizotypal Personality Questionnaire. *Personality Disorders: Theory, Research, and Treatment*, *5*(4), 397.

Gumley, A., & Schwannauer, M. (2006). *Staying well after psychosis: A cognitive interpersonal approach to recovery and relapse prevention*: John Wiley & Sons.

Ham, S., Kim, T. K., Chung, S., & Im, H.-I. (2017). Drug abuse and psychosis: new insights into drug-induced psychosis. *Experimental neurobiology, 26*(1), 11-24.

Hamm, J. A., Leonhardt, B. L., Ridenour, J., Lysaker, J. T., & Lysaker, P. H. (2018). Phenomenological and recovery models of the subjective experience of psychosis: discrepancies and implications for treatment. *Psychosis, 10*(4), 340-350.

Hans, S. L., Auerbach, J. G., Nuechterlein, K. H., Asarnow, R. F., Asarnow, J., Styr, B., & Marcus, J. (2009). Neurodevelopmental factors associated with schizotypal symptoms among adolescents at risk for schizophrenia. *Development and psychopathology*, *21*(4), 1195-1210.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129-154.

Harrington, L., Siegert, R., & McClure, J. (2005). Theory of mind in schizophrenia: a critical review. *Cognitive neuropsychiatry*, *10*(4), 249-286.

Hart, J. R., Venta, A., & Sharp, C. (2017). Attachment and thought problems in an adolescent inpatient sample: The mediational role of theory of mind. *Comprehensive psychiatry, 78*, 38-47.

Hasson-Ohayon, I., Cheli, S., & Lysaker, P. H. (2021). Emerging psychotherapeutic approaches to addressing self-experience in schizophrenia spectrum disorders.

Hasson-Ohayon, I., Kravetz, S., & Lysaker, P. H. (2017). The special challenges of

    psychotherapy with persons with psychosis: Intersubjective metacognitive model of

    agreement and shared meaning. *Clinical psychology & psychotherapy, 24*(2), 428-

    440.

Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis

    (2nd ed.). New York: Guilford Press.

Henquet, C., Krabbendam, L., Dautzenberg, J., Jolles, J., & Merckelbach, H. (2005).

    Confusing thoughts and speech: source monitoring and psychosis. *Psychiatry*

    research, 133(1), 57-63.

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: a

    quantitative review of the evidence. *Neuropsychology*, *12*(3), 426.

Hommes, J., Krabbendam, L., Versmissen, D., Kircher, T., van Os, J., & van Winkel, R.

    (2012). Self-monitoring as a familial vulnerability marker for psychosis: an analysis

    of patients, unaffected siblings and healthy controls. *Psychol Med*, *42*, 235-245.

Honig, A., Romme, M. A., Ensink, B. J., Escher, S. D., Pennings, M. H., & DEVRIES, M.

    W. (1998). Auditory hallucinations: a comparison between patients and

    nonpatients. *The Journal of nervous and mental disease*, *186*(10), 646-651.

Humpston, C.S., Linden, D.E., & Evans, L.H. (2017). Deficits in reality and internal source

    monitoring of actions are associated with the positive dimension of schizotypy.

    Psychiatry research, 250, 44-49.

Itaguchi, Y., Sugimori, E., & Fukuzawa, K. (2018). Schizotypal traits and forearm motor control against self-other produced action in a bimanual unloading task. *Neuropsychologia*, *113*, 43-51.

Johns, L., Allen, P., Valli, I., Winton-Brown, T., Broome, M., Woolley, J., Tabraham, P., Day, F., Howes, O., & Wykes, T. (2010). Impaired verbal self-monitoring in individuals at high risk of psychosis. *Psychological medicine, 40*(9), 1433-1442.

Johns, L.C., Gregg, L., Allen, P., & McGUIRE, P.K. (2006). Impaired verbal self-monitoring in psychosis: effects of state, trait and diagnosis. *Psychological medicine, 36*(4), 465-474.

Johns, L. C., & Van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical psychology review*, *21*(8), 1125-1141.

Johnson, M.K. (2006). Memory and reality. *American Pysychologist*, *61*(8), 760.

Johnson, M.K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology*, *117*(4), 390.

Johnson, M. K., Nolde, S. F., & De Leonardis, D. M. (1996). Emotional focus and source monitoring. *Journal of Memory and Language*, *35*(2), 135-156.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3.

Johnson, M. K., & Suengas, A. G. (1989). Reality monitoring judgments of other people's memories. *Bulletin of the Psychonomic Society*, *27*(2), 107-110.

Johnstone, E. C., Ebmeier, K. P., Miller, P., Owens, D. G., & Lawrie, S. M. (2005). Predicting schizophrenia: findings from the Edinburgh high-risk study. *The British Journal of Psychiatry*, *186*(1), 18-25.

Kantrowitz, J. T., Hoptman, M. J., Leitman, D. I., Silipo, G., & Javitt, D. C. (2014). The 5% difference: early sensory processing predicts sarcasm perception in schizophrenia and schizoaffective disorder. *Psychological medicine*, *44*(1).

Keefe, R.S., Arnold, M., Bayen, U., & Harvey, P. (1999). Source monitoring deficits in patients with schizophrenia; a multinomial modelling analysis. *Psychological medicine, 29*(4), 903-914.

Kendler, K. S., Gallagher, T. J., Abelson, J. M., & Kessler, R. C. (1996). Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a US community sample: the National Comorbidity Survey. *Archives of general psychiatry*, *53*(11), 1022-1031.

Kim, K. R., Lee, S. Y., Kang, J. I., Kim, B. R., Choi, S. H., Park, J. Y., Lee, E., An, S. K., & Kwon, J. S. (2011). Clinical efficacy of individual cognitive therapy in reducing psychiatric symptoms in people at ultra-high risk for psychosis. *Early intervention in psychiatry, 5*(2), 174-178.

Kinderman, P., & Bentall, R. P. (1997). Causal attributions in paranoia and depression: internal, personal, and situational attributions for negative events. *Journal of abnormal psychology*, *106*(2), 341.

Klagsbrun, M., & Bowlby, J. (1976). Responses to separation from parents: a clinical test for young children. *British Journal of Projective Psychology & Personality Study*.

Knoblich, G., Stottmeister, F., & Kircher, T. (2004). Self-monitoring in patients with schizophrenia. *Psychological medicine*, *34*(8), 1561.

Koreki, A., Funayama, M., Terasawa, Y., Onaya, M., & Mimura, M. (2021). Aberrant interoceptive accuracy in patients with schizophrenia performing a heartbeat counting task. *Schizophrenia Bulletin Open*, *2*(1), sgaa067.

Korver-Nieberg, N., Fett, A. K. J., Meijer, C. J., Koeter, M. W., Shergill, S. S., de Haan, L., & Krabbendam, L. (2013). Theory of mind, insecure attachment and paranoia in adolescents with early psychosis and healthy controls. *Australian & New Zealand Journal of Psychiatry*, *47*(8), 737-745.

Kwapil, T. R., & Barrantes-Vidal, N. (2015). Schizotypy: looking back and moving forward. *Schizophrenia bulletin*, *41*(suppl_2), S366-S373.

Kwapil, T. R., Gross, G. M., Silvia, P. J., & Barrantes-Vidal, N. (2013). Prediction of psychopathology and functional impairment by positive and negative schizotypy in the Chapmans' ten-year longitudinal study. *Journal of abnormal psychology, 122*(3), 807.

Lagioia, A., Eliez, S., Schneider, M., Simons, J.S., Van der Linden, M., & Debbané, M. (2011). Neural correlates of reality monitoring during adolescence. *NeuroImage, 55*(3), 1393-1400.

Langdon, R., & Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. *Cognition, 71*(1), 43-71.

Langdon, R., & Brock, J. (2008). Hypo-or hyper-mentalizing: It all depends upon what one means by "mentalizing". *Behavioral and Brain Sciences*, *31*(3), 274-275.

Larøi, F., Van der Linden, M., & Marczewski, P. (2004). The effects of emotional salience, cognitive effort and meta-cognitive beliefs on a reality monitoring task in hallucination-prone subjects. *British Journal of Clinical Psychology, 43*(3), 221-233.

Launay, G., & Slade, P. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences*, *2*(3), 221-234.

Lemaitre, A.-L., Luyat, M., Lafargue, G. (2016). Individuals with pronounced schizotypal traits are particularly successful in tickling themselves. *Consciousness and Cognition*, *41*, 64-71.

Lenzenweger, M. F. (2015). Thinking clearly about schizotypy: hewing to the schizophrenia liability core, considering interesting tangents, and avoiding conceptual quicksand. *Schizophrenia Bulletin*, *41*(suppl_2), S483-S491.

Li, E. T., Carracher, E., & Bird, T. (2020). Linking childhood emotional abuse and adult depressive symptoms: the role of mentalizing incapacity. *Child abuse & neglect*, *99*, 104253.

Luyten, P., van Houdenhove, B., Lemma, A., Target, M., & Fonagy, P. (2012). A mentalization-based approach to the understanding and treatment of functional somatic disorders. *Psychoanalytic Psychotherapy, 26*(2), 121-140.

Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacci, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia: associations with neurocognition, symptoms, insight and quality of life. *Acta psychiatrica scandinavica*, *112*(1), 64-71.

Lysaker, P. H., Olesek, K. L., Warman, D. M., Martin, J. M., Salzman, A. K., Nicolò, G., ... & Dimaggio, G. (2011). Metacognition in schizophrenia: Correlates and stability of deficits in theory of mind and self-reflectivity. *Psychiatry Research*, *190*(1), 18-22.

Lysaker, P. H., Hamm, J. A., Vohs, J., Kukla, M., Pattison, M. L., Leonhardt, B. L., & Lysaker, J. T. (2018). Understanding the Course of Self-disorders and Alterations in

Self-Experience in Schizophrenia: Implications from Research on Metacognition. *Current Psychiatry Reviews, 14*(3), 160-170.

Lysaker, P. H., Gagen, E., Klion, R., Zalzala, A., Vohs, J., Faith, L. A., ... & Hasson-Ohayon, I. (2020). Metacognitive reflection and insight therapy: a recovery-oriented treatment approach for psychosis. *Psychology research and behavior management*, *13*, 331.

Lysaker, P. H., Cheli, S., Dimaggio, G., Buck, B., Bonfils, K. A., Huling, K., ... & Lysaker, J. T. (2021). Metacognition, social cognition, and mentalizing in psychosis: are these distinct constructs when it comes to subjective experience or are we just splitting hairs?. *BMC psychiatry*, *21*(1), 1-14.

MacBeth, A., Schwannauer, M., & Gumley, A. (2008). The association between attachment style, social mentalities, and paranoid ideation: An analogue study. *Psychology and Psychotherapy: Theory, Research and Practice*, *81*(1), 79-93.

Mason, O. (2015). The assessment of schizotypy and its clinical relevance. *Schizophrenia bulletin*, *41*(suppl_2), S374-S385.

Mason, O., Startup, M., Halpin, S., Schall, U., Conrad, A., & Carr, V. (2004). Risk factors for transition to first episode psychosis among individuals with 'at-risk mental states'. *Schizophrenia research*, *71*(2-3), 227-237.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual differences*, *18*(1), 7-13.

Markham, R. (1991). Development of reality monitoring for performed and imagined actions. *Perceptual and Motor Skills*, *72*(3_suppl), 1347-1354

Mazza, M., De Risio, A., Surian, L., Roncone, R., & Casacchia, M. (2001). Selective impairments of theory of mind in people with schizophrenia. *Schizophrenia research*, *47*(2-3), 299-308

McGlashan, T. H., Miller, T. J., Woods, S. W., Hoffman, R. E., & Davidson, L. (2001). Instrument for the assessment of prodromal symptoms and states. In *Early intervention in psychotic disorders* (pp. 135-149). Springer, Dordrecht.

McGorry, P. D., Purcell, R., Hickie, I. B., Yung, A. R., Pantelis, C., & Jackson, H. J. (2007). Clinical staging: a heuristic model for psychiatry and youth mental health. *Medical Journal of Australia*, *187*(S7), S40-S42.

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiologic reviews*, *30*(1), 67-76.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American psychologist*, *17*(12), 827.

Michail, M., & Birchwood, M. (2009). Social anxiety disorder in first-episode psychosis: incidence, phenomenology and relationship with paranoia. *The British Journal of Psychiatry*, *195*(3), 234-241.

Miettunen, J., Veijola, J., Isohanni, M., Paunio, T., Freimer, N., Jääskeläinen, E., ... & Lichtermann, D. (2011). Identifying schizophrenia and other psychoses with psychological scales in the general population. *The Journal of nervous and mental disease*, *199*(4), 230-238.

Moeller, F.G., Barratt, E.S., Dougherty, D.M., Schmitz, J.M., & Swann, A.C. (2001). Psychiatric aspects of impulsivity. *American Journal of Psychiatry, 158*(11), 1783-1793.

Montag, C., Dziobek, I., Richter, I. S., Neuhaus, K., Lehmann, A., Sylla, R., ... & Gallinat, J. (2011). Different aspects of theory of mind in paranoid schizophrenia: evidence from a video-based assessment. *Psychiatry research*, *186*(2-3), 203-209.

Morrison, A. P., & Haddock, G. (1997). Cognitive factors in source monitoring and auditory hallucinations. *Psychological Medicine*, *27*(3), 669-679.

Morosan, L., Badoud, D., Salaminios, G., Eliez, S., Van der Linden, M., Heller, P., & Debbané, M. (2018). Patterns of source monitoring bias in incarcerated youths with and without conduct problems. *Cognitive Neuropsychiatry, 23*(1), 15-27.

Müller, S., Wendt, L. P., Spitzer, C., Masuhr, O., Back, S. N., & Zimmermann, J. (2020). A Critical Evaluation of the Reflective Functioning Questionnaire.

Müller, V.I., Langner, R., Cieslik, E.C., Rottschy, C., Eickhoff, S.B. (2015). Interindividual differences in cognitive flexibility: influence of gray matter volume, functional connectivity and trait impulsivity. *Brain Structure and Function*, *220*(4), 2401-2414.

Murphy, E. R., Brent, B. K., Benton, M., Pruitt, P., Diwadkar, V., Rajarethinam, R. P., & Keshavan, M. S. (2010). Differential processing of metacognitive evaluation and the neural circuitry of the self and others in schizophrenia: a pilot study. *Schizophrenia research*, *116*(2-3), 252-258.

Murray, R., van Os, J., McKenzie, K., Allardyce, J., Goel, R., McCreadie, R., & Boydell, J. (2003) The incidence of schizophrenia in ethnic minorities in London—Evidence for social aetiological factors. *Schizophrenia Research*, *1*(60), 47.

Murre, J.M., Janssen, S.M., Rouw, R., & Meeter, M.J. (2013). The rise and fall of immediate and delayed memory for verbal and visuospatial information from late childhood to late adulthood. Acta Psychologica, *142*(1), 96-107.

Nelson, B., Yuen, H. P., Lin, A., Wood, S. J., McGorry, P. D., Hartmann, J. A., & Yung, A. R. (2016). Further examination of the reducing transition rate in ultra high risk for psychosis samples: the possible role of earlier intervention. *Schizophrenia Research*, *174*(1-3), 43-49.

Nelson, M. T., Seal, M. L., Pantelis, C., & Phillips, L. J. (2013). Evidence of a dimensional relationship between schizotypy and schizophrenia: a systematic review. *Neuroscience & Biobehavioral Reviews*, *37*(3), 317-327.

Nuzzo, R.L. (2017), An Introduction to Bayesian Data Analysis for Correlations. PM&R, 9: 1278-1282.

Oezgen, M., & Grant, P. (2018). Odd and disorganized–comparing the factor structure of the three major schizotypy inventories. *Psychiatry Research*, *267*, 289-295.

Ng, R., Fish, S., & Granholm, E. (2015). Insight and theory of mind in schizophrenia. *Psychiatry research*, *225*(1-2), 169-174.

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence?. *Nature reviews neuroscience*, *9*(12), 947-957.

Pedersen, A., & Rist, F. (2001). Implicit memory in schizotypal subjects and normal controls: effects of a secondary task on sequence learning. *Perceptual and motor skills*, *92*(2), 349-367.

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. Delusions Inventory). *Schizophrenia bulletin*, *25*(3), 553-576.

Peters, M.J., Smeets, T., Giesbrecht, T., Jelicic, M., & Merckelbach, H. (2007). Confusing action and imagination: action source monitoring in individuals with schizotypal traits. *The Journal of nervous and mental disease, 195*(9), 752-757.

Peters, E., Ward, T., Jackson, M., Morgan, C., Charalambides, M., McGuire, P., Woodruff, P., Jacobsen, P., Chadwick, P., & Garety, P. A. (2016). Clinical, socio-demographic

and psychological characteristics in individuals with persistent psychotic experiences with and without a "need for care". *World psychiatry, 15*(1), 41-52.

Pickering, L., Simpson, J., & Bentall, R. P. (2008). Insecure attachment predicts proneness to paranoia but not hallucinations. *Personality and Individual Differences*, *44*(5), 1212-1224.

Pickup, G. (2006). Theory of mind and its relation to schizotypy. *Cognitive Neuropsychiatry, 11*(2), 117-192.

Pickup, G. J., & D FRITH, C. (2001). Theory of mind impairments in schizophrenia: symptomatology, severity and specificity. *Psychological medicine*, *31*(2), 207.

Piskulic, D., Liu, L., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., McGlashan, T. H., Perkins, D. O., Seidman, L. J., Tsuang, M. T., & Walker, E. F. (2016). Social cognition over time in individuals at clinical high risk for psychosis: findings from the NAPLS-2 cohort. *Schizophrenia Research, 171*(1-3), 176-181.

Pontillo, M., Guerrera, S., Santonastaso, O., Tata, M. C., Averna, R., Vicari, S., & Armando, M. (2017). An overview of recent findings on social anxiety disorder in adolescents and young adults at clinical high risk for psychosis. *Brain sciences*, *7*(10), 127.

Poulton, R., Caspi, A., Moffitt, T. E., Cannon, M., Murray, R., & Harrington, H. (2000). Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. *Archives of general psychiatry*, *57*(11), 1053-1058.

Postmes, L., Sno, H., Goedhart, S., Van Der Stel, J., Heering, H., & De Haan, L. (2014). Schizophrenia as a self-disorder due to perceptual incoherence. *Schizophrenia Research, 152*(1), 41-50.

Poznyak, E., Morosan, L., Perroud, N., Speranza, M., Badoud, D., & Debbané, M. (2019). Roles of age, gender and psychological difficulties in adolescent mentalizing. *Journal of adolescence*, *74*, 120-129.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, *1*(4), 515-526.

Price, C.J., & Devlin, J.T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246-253.

Raballo, A., & Larøi, F. (2009). Clinical staging: a new scenario for the treatment of psychosis. *The Lancet*, *374*(9687), 365-367.

Raballo, A., Nelson, B., Thompson, A., & Yung, A. (2011). The comprehensive assessment of at-risk mental states: from mapping the onset to mapping the structure. *Schizophrenia research*, *127*(1-3), 107-114.

Raine, A. (2006). Schizotypal personality: neurodevelopmental and psychosocial trajectories. *Annu. Rev. Clin. Psychol., 2*, 291-326.

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia bulletin*, *17*(4), 555-564.

Rietdijk, J., Ising, H. K., Dragt, S., Klaassen, R., Nieman, D., Wunderink, L., van der Gaag, M. (2013). Depression and social anxiety in help-seeking patients with an ultra-high risk for developing psychosis. Psychiatry Research, 209(3), 309–313.

Romero-Garcia, R., Seidlitz, J., Whitaker, K. J., Morgan, S. E., Fonagy, P., Dolan, R. J., ... & Bullmore, E. T. (2020). Schizotypy-related magnetization of cortex in healthy adolescence is colocated with expression of schizophrenia-related genes. *Biological psychiatry*, *88*(3), 248-259.

Rossouw, T. I., & Fonagy, P. (2012). Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(12), 1304-1313. e1303.

Rudrauf D. Structure-function relationships behind the phenomenon of cognitive resilience in neurology: insights for neuroscience and medicine. Adv Neurosci 2014;2014:1-28.

Ruhrmann, S., Schultze-Lutter, F., Salokangas, R. K., Heinimaa, M., Linszen, D., Dingemans, P., ... & Klosterkötter, J. (2010). Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. *Archives of general psychiatry*, *67*(3), 241-251.

Sahakyan, L., & Kwapil, T.R. (2016). Positive schizotypy and negative schizotypy are associated with differential patterns of episodic memory impairment. *Schizophrenia Research: Cognition*, *5*, 35-40.

Salaminios, G. and Debbané, M. (2021) A Mentalization-Based Treatment Framework to Support the Recovery of the Self in Emerging Psychosis During Adolescence. In Hasson-Ohayon, I. and Lysaker, P. (Eds.) *The Recovery of the Self in Psychosis: Contributions from Metacognitive and Mentalization Based Oriented Psychotherapy.* Routledge

Salaminios, G., Morosan, L., Toffel, E., Tanzer, M., Eliez, S., Badoud, D., ... & Debbané, M. (2020a). Associations between schizotypal personality features, mentalizing difficulties and thought problems in a sample of community adolescents. *Early Intervention in Psychiatry*.

Salaminios, G., Morosan, L., Toffel, E., Tanzer, M., Eliez, S., & Debbané, M. (2020b). Self-Monitoring for speech and its links to age, cognitive effort, schizotypal trait

expression and impulsivity during adolescence. *Cognitive neuropsychiatry*, *25*(3), 215-230.

Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia bulletin, 29*(3), 427-444.

Sarfati, Y., Passerieux, C., & Hardy-Baylé, M. C. (2000). Can verbalization remedy the theory of mind deficit in schizophrenia?. *Psychopathology*, *33*(5), 246-251.

Sarfati, Y., Hardy-Baylé, M. C., Brunet, E., & Widlöcher, D. (1999). Investigating theory of mind in schizophrenia: influence of verbalization in disorganized and non disorganized patients. *Schizophrenia research*, *37*(2), 183-190.

Sarfati, Y., Hardy-Baylé, M. C., Besche, C., & Widlöcher, D. (1997). Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophrenia   research*, *25*(3), 199-209.

Schultze-Lutter, F., Debbané, M., Theodoridou, A., Wood, S. J., Raballo, A., Michel, C., ... & Uhlhaas, P. J. (2016). Revisiting the basic symptom concept: toward translating risk symptoms for psychosis into neurobiological targets. *Frontiers in psychiatry*, *7*, 9.

Schultze-Lutter, F., Klosterkötter, J., & Ruhrmann, S. (2014). Improving the clinical prediction of   psychosis by combining ultra-high risk criteria and cognitive basic symptoms. *Schizophrenia research*, *154*(1-3), 100-106.

Sergi, M. J., Rassovsky, Y., Widmark, C., Reist, C., Erhart, S., Braff, D. L., Marder, S. R., & Green, M. F. (2007). Social cognition in schizophrenia: relationships with neurocognition and negative symptoms. *Schizophrenia Research, 90*(1-3), 316-324.

Shad, M. U., Brent, B. K., & Keshavan, M. S. (2011). Neurobiology of self-awareness deficits in schizophrenia: A hypothetical model. *Asian journal of psychiatry*, *4*(4), 248-254.

Shah, J., Eack, S. M., Montrose, D. M., Tandon, N., Miewald, J. M., Prasad, K. M., & Keshavan, M. S. (2012). Multivariate prediction of emerging psychosis in adolescents at high risk for schizophrenia. *Schizophrenia Research*, *141*(2-3), 189-196.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia*, *45*(13), 3054-3067.

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology*, *18*(1), 55-67.

Sharp, C., & Fonagy, P. (2008). The parent's capacity to treat the child as a psychological agent: Constructs, measures and implications for developmental psychopathology. *Social development, 17*(3), 737-754.

Sharp, C., Ha, C., Carbone, C., Kim, S., Perry, K., Williams, L., & Fonagy, P. (2013). Hypermentalizing in adolescent inpatients: treatment effects and association with borderline traits. *Journal of personality disorders*, *27*(1), 3-18.

Sharp, C., Ha, C., Carbone, C., Kim, S., Perry, K., Williams, L., & Fonagy, P. (2013). Hypermentalizing in adolescent inpatients: treatment effects and association with borderline traits. *Journal of personality disorders*, *27*(1), 3-18.

Sharpley, M. S., & Peters, E. R. (1999). Ethnicity, class and schizotypy. *Social psychiatry and psychiatric epidemiology*, *34*(10), 507-512.

Shaw, C., Lo, C., Lanceley, A., Hales, S., & Rodin, G. (2020). The assessment of mentalization: measures for the patient, the therapist and the interaction. *Journal of Contemporary Psychotherapy*, *50*(1), 57-65.

Sheinbaum, T., Bedoya, E., Ros-Morente, A., Kwapil, T. R., & Barrantes-Vidal, N. (2013).

　　Association between attachment prototypes and schizotypy dimensions in two

　　independent non-clinical samples of Spanish and American young adults. *Psychiatry*

　　research, *210*(2), 408-413.

Simon, A. E., Velthorst, E., Nieman, D. H., Linszen, D., Umbricht, D., & de Haan, L. (2011).

　　Ultra high-risk state for psychosis and non-transition: a systematic

　　review. *Schizophrenia research*, *132*(1), 8-17.

Sommer, I. E., Derwort, A. M., Daalman, K., de Weijer, A. D., Liddle, P. F., & Boks, M. P.

　　(2010). Formal thought disorder in non-clinical individuals with auditory verbal

　　hallucinations. *Schizophrenia research*, *118*(1-3), 140-145.

Simeonova, D. I., Nguyen, T., & Walker, E. F. (2014). Psychosis risk screening in clinical

　　high-risk adolescents: a longitudinal investigation using the Child Behavior Checklist.

　　Schizophrenia Research, 159(1), 7-13.

Simons, J. S., Davis, S. W., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2006).

　　Discriminating imagined from perceived information engages brain areas implicated

　　in schizophrenia. *Neuroimage*, *32*(2), 696-703.

Spark, J., Gawęda, Ł., Allott, K., Hartmann, J. A., Jack, B. N., Koren, D., ... & Nelson, B.

　　Distinguishing schizophrenia spectrum from non-spectrum disorders among young

　　patients with first episode psychosis and at high clinical risk: The role of basic self-

　　disturbance and neurocognition. *Schizophrenia Research*, *228*, 19-28.

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Van Engeland, H. (2007). Theory of mind in

　　schizophrenia: meta-analysis. *The British Journal of Psychiatry, 191*(1), 5-13.

Steele, H., Steele, M., & Croft, C. (2008). Early attachment predicts emotion recognition at 6

　　and 11 years old. *Attachment & Human Development*, *10*(4), 379-393.

Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K.,
Stefanis, C. N., ... & Van Os, J. (2002). Evidence that three dimensions of psychosis
have a distribution in the general population. *Psychological medicine*, *32*(2), 347-358.

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual review of
psychology*, *52*(1), 83-110.

Stephane, M. (2019). The self, agency and spatial externalizations of inner verbal thoughts,
and auditory verbal hallucinations. *Frontiers in psychiatry*, *10*, 668.

Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and
developmental psychology*: Karnac Books.

Stevens, J., Schmied, V., Burns, E., & Dahlen, H. G. (2019). Skin-to-skin contact and what
women want in the first hours after a caesarean section. *Midwifery*, *74*, 140-146.

Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for
perceived and imagined complex events. *Journal of Experimental Psychology:
General*, *117*(4), 377.

Sugimori, E., Asai, T., Tanno, Y. (2011). Sense of agency over thought: external
misattribution of thought in a memory task and proneness to auditory hallucination.
Consciousness and Cognition, *20*(3), 688-695.

Sussman, A.L. (2001). Reality monitoring of performed and imagined interactive events:
Developmental and contextual effects. *Journal of Experimental Child Psychology,
79*(2), 115-138.

Tanzer, M., Salaminios, G., Morosan, L., Campbell, C., & Debbané, M. (2020). Self-blame
mediates the link between childhood neglect experiences and internalizing symptoms
in low-risk adolescents. *Journal of Child & Adolescent Trauma*, 1-11.

Taubner, S., Kessler, H., Buchheim, A., Kächele, H., & Staun, L. (2011). The role of mentalization in the psychoanalytic treatment of chronic depression. *Psychiatry: Interpersonal & Biological Processes, 74*(1), 49-57.

Taylor, J., Rastle, K., & Davis, M.H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychological Bulletin, *139*(4), 766.

Thompson, A. D., Bartholomeusz, C., & Yung, A. R. (2011). Social cognition deficits and the 'ultra high risk' for psychosis population: a review of literature. *Early intervention in psychiatry*, *5*(3), 192-202.

Thompson, A., Papas, A., Bartholomeusz, C., Allott, K., Amminger, G. P., Nelson, B., ... & Yung, A. (2012). Social cognition in clinical "at risk" for psychosis and first episode psychosis populations. *Schizophrenia research*, *141*(2-3), 204-209.

Tien, A. Y. (1991). Distribution of hallucinations in the population. *Social psychiatry and psychiatric epidemiology*, *26*(6), 287-292.

Tiliopoulos, N., & Goodall, K. (2009). The neglected link between adult attachment and schizotypal personality traits. *Personality and Individual Differences*, *47*(4), 299-304.

Tsakiris, M., Carpenter, L., James, D., & Fotopoulou, A. (2010). Hands only illusion: multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research, 204*(3), 343-352.

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in cognitive sciences, 11*(4), 153-157.

van den Bergh, D., Clyde, M. A., Gupta, A. R. K. N., de Jong, T., Gronau, Q. F., Marsman, M., & Wagenmakers, E. J. (2021). A tutorial on Bayesian multi-model linear regression with BAS and JASP. *Behavior research methods*, 1-21.

van Os, J., Kenis, G., & Rutten, B. P. (2010). The environment and schizophrenia. *Nature, 468*(7321), 203.

Van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. J. P. M. (2009). A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychological medicine*, *39*(2), 179.

Vanderwal, T., Hunyadi, E., Grupe, D. W., Connors, C. M., & Schultz, R. T. (2008). Self, mother and abstract other: an fMRI study of reflective social processing. *Neuroimage, 41*(4), 1437-1446.

Varese, F., & Bentall, R. P. (2011). The metacognitive beliefs account of hallucinatory experiences: a literature review and meta-analysis. *Clinical psychology review*, *31*(5), 850-864.

Verdoux, H., & van Os, J. (2002). Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophrenia research*, *54*(1-2), 59-65.

Versmissen, D., Janssen, I., Myin-Germeys, I., Mengelers, R., a Campo, J., van Os, J., & Krabbendam, L. (2008). Evidence for a relationship between mentalising deficits and paranoia over the psychosis continuum. *Schizophrenia Research*, *99*(1-3), 103-110.

Versmissen, D., Myin-Germeys, I., Janssen, I., Franck, N., Georgieff, N., Campo, J., ... & Krabbendam, L. (2007). Impairment of self-monitoring: part of the endophenotypic risk for psychosis. *The British Journal of Psychiatry*, *191*(S51), s58-s62.

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology, 38*(2), 114-125.

Vinogradov, S., Luks, T. L., Schulman, B. J., & Simpson, G. V. (2008). Deficit in a neural correlate of reality monitoring in schizophrenia patients. *Cerebral Cortex*, *18*(11), 2532-2539.

Vinogradov, S., Luks, T. L., Simpson, G. V., Schulman, B. J., Glenn, S., & Wong, A. E. (2006). Brain activation patterns during memory of cognitive agency. *Neuroimage*, *31*(2), 896-905.

Wechsler, D. (1955). Manual for the Wechsler adult intelligence scale. The Psychological Corporation. *New York*.

Weijers, J., Ten Kate, C., Viechtbauer, W., Rampaart, L. J. A., Eurelings, E. H. M., & Selten, J. P. (2020). Mentalization-based treatment for psychotic disorder: a rater-blinded, multi-center, randomized controlled trial. *Psychological Medicine*, 1-10.

Weijers, J., Fonagy, P., Eurelings-Bontekoe, E., Termorshuizen, F., Viechtbauer, W., & Selten, J. (2018). Mentalizing impairment as a mediator between reported childhood abuse and outcome in nonaffective psychotic disorder. Psychiatry Research, 259, 463–469.

Weijers, J., ten Kate, C., Eurelings-Bontekoe, E., Viechtbauer, W., Rampaart, R., Bateman, A., & Selten, J.-P. (2016). Mentalization-based treatment for psychotic disorder: protocol of a randomized controlled trial. *BMC psychiatry, 16*(1), 191

Weiss, S. J., Wilson, P., & Morrison, D. (2004). Maternal tactile stimulation and the neurodevelopment of low birth weight infants. *Infancy*, *5*(1), 85-107.

Whiteside, S.P., & Lynam, D.R. (2001). The five factor model and impulsivity: Using a

 structural model of personality to understand impulsivity. *Personality and Individual*

 Differences, 30(4), 669-689.

Wilbertz, T., Deserno, L., Horstmann, A., Neumann, J., Villringer, A., Heinze, H.-J.,

 Boehler, C.N., & Schlagenhauf, F. (2014). Response inhibition and its relation to

 multidimensional impulsivity. *Neuroimage*, *103*, 241-248.

Young, K. S., Sandman, C. F., & Craske, M. G. (2019). Positive and negative emotion

 regulation in adolescence: links to anxiety and depression. *Brain sciences*, *9*(4), 76.

Yung, A. R., Phillips, L. J., McGorry, P. D., McFarlane, C. A., Francey, S., Harrigan, S., ... &

 Jackson, H. J. (1998). Prediction of psychosis. *The British Journal of*

 Psychiatry, *172*(S33), 14-20.

Yung, A. R., McGorry, P. D., McFarlane, C. A., Jackson, H. J., Patton, G. C., & Rakkar, A.

 (1996). Monitoring and Care of Young People at Incipient Risk of Psychosis.

 Schizophrenia Bulletin, 22(2), 283–303. https://doi.org/10.1093/schbul/22.2.283

Appendices

Appendix 1: Items of the Schizotypal Personality Questionnaire (SPQ) (English version, Raine, 1991)

Interpersonal Scale

Excessive Social Anxiety

2. I sometimes avoid going to places where there will be many people because I will get anxious.

11. I get very nervous when I have to make polite conversation.

20. Do you ever get nervous when someone is walking behind you?

29. I get anxious when meeting people for the first time.

38. Do you often feel nervous when you are in a group of unfamiliar people?

46. I feel very uncomfortable in social situations involving unfamiliar people.

54. I would feel very anxious if I had to give a speech in front of a large group of people.

71. I feel very uneasy talking to people I do not know well.

No Close Friends

6. I have little interest in getting to know other people.

15. I prefer to keep myself to myself.

24. I am mostly quiet when with other people.

33. I find It hard to be emotionally close to other people.

41. Do you feel that there Is no one you are really close to outside of your immediate family, or people you can confide In or talk to about personal problems?

49. Writing letters to friends is more trouble than it is worth.

57. I tend to keep in the background on social occasions.

62. I attach little Importance to having close friends.

66. Do you feel that you cannot get "close" to people?

Constricted Affect

8. People sometimes find me aloof and distant.

17. I am not good at expressing my true feelings by the way I talk and look.

26. I rarely laugh and smile.

35. My "nonverbal" communication (smiling and nodding during a conversation) Is not very good.

43. I am poor at returning social courtesies and gestures.

51. I tend to avoid eye contact when conversing with others.

68. I do not have an expressive and lively way of speaking.

73. I tend to keep my feelings to myself.

Cognitive-perceptual scale

Ideas of Reference

1. Do you sometimes feel that things you see on the TV or read In the newspaper have a special meaning for you?

10. I am aware that people notice me when I go out for a meal or to see a film.

19. Do some people drop hints about you or say things with a double meaning?

28. Have you ever noticed a common event or object that seemed to be a special sign for you?

37. Do you sometimes see special meanings in advertisements, shop windows, or In the way things are arranged around you?

45. When shopping do you get the feeling that other people are taking notice of you?

53. When you see people talking to each other, do you often wonder If they are talking about you?

60. Do you sometimes feel that other people are watching you?

63. Do you sometimes feel that people are talking about you?

Odd Beliefs or Magical Thinking

3. Have you had experiences with the supernatural?

12. Do you believe in telepathy (mind-reading)?

21. Are you sometimes sure that other people can tell what you are thinking?

30. Do you believe in clairvoyancy (psychic forces, fortune telling)?

39. Can other people feel your feelings when they are not there?

47. Have you had experiences with astrology, seeing the future, UFOs, ESP, or a sixth sense?

55. Have you ever felt that you are communicating with another person telepathically (by mind-reading)?

Unusual Perceptual Experiences

4. Have you often mistaken objects or shadows for people, or noises for voices?

13. Have you ever had the sense that some person or force is around you, even though you cannot see anyone?

22. When you look at a person, or yourself in a mirror, have you ever seen the face change right before your eyes?

31. I often hear a voice speaking my thoughts aloud.

40. Have you ever seen things invisible to other people?

48. Do everyday things seem unusually large or small?

56. Does your sense of smell sometimes become unusually strong?

61. Do you ever suddenly feel distracted by distant sounds that you are not normally aware of?

64. Are your thoughts sometimes so strong that you can almost hear them?

Suspiciousness

9. I am sure I am being talked about behind my back.

18. Do you often feel that other people have it in for you?

27. Do you sometimes get concerned that friends or coworkers are not really loyal or trustworthy?

36. I feel I have to be on my guard even with friends.

44. Do you often pick up hidden threats or put-downs from what people say or do?

52. Have you found that It is best not to let other people know too much about you?

59. I often feel that others have it In for me.

65. Do you often have to keep an eye out to stop people from taking advantage of you?

Disorganization scale

Odd or Eccentric Behavior

5. Other people see me as slightly eccentric (odd).

14. People sometimes comment on my unusual mannerisms and habits.

23. Sometimes other people think that I am a little strange.

32. Some people think that I am a very bizarre person.

67. I am an odd, unusual person.

70. I have some eccentric (odd) habits.

74. People sometimes stare at me because of my odd appearance.

Odd Speech

7. People sometimes find it hard to understand what I am saying.

16. I sometimes jump quickly from one topic to another when speaking.

25. I sometimes forget what I am trying to say.

34. I often ramble on too much when speaking.

42. Some people find me a bit vague and elusive during a conversation.

50. I sometimes use words In unusual ways.

58. Do you tend to wander off the topic when having a conversation?

69. I find it hard to communicate clearly what I want to say to people.

72. People occasionally comment that my conversation is confusing.

Appendix 2: The Reflective Functioning Questionnaire (RFQ-8) (English version, Fonagy et al., 2016)

Please work through the next 8 statements. For each statement, choose a number between 1 and 7 to say how much you disagree or agree with the statement, and write it beside the statement. Do not think too much about it – your initial responses are usually the best. Thank you.

Use the following scale from 1 to 7:

Strongly disagree	1	2	3	4	5	6	7	Strongly agree

1. __ People's thoughts are a mystery to me

2. __ I don't always know why I do what I do

3. __ When I get angry I say things without really knowing why I am saying them

4. __ When I get angry I say things that I later regret

5. __ If I feel insecure I can behave in ways that put others' backs up

6. __ Sometimes I do things without really knowing why

7. __ I always know what I feel

8. __ Strong feelings often cloud my thinking

Appendix 3: Items of the Youth and Adult Self Report (YSR/ASR) Thought Problems Scale (English version, Achenbach, 1991)

YSR Thought Problems Scale items	ASR Thought Problems Scale items
9: I can't get my mind off certain thoughts	*9:* I can't get my mind off certain thoughts
18: I deliberately try to hurt or kill myself	*18:* I deliberately try to hurt or kill myself
40: I hear sounds or voices that other people think aren't there	*40:* I hear sounds or voices that other people think aren't there
46: Parts of my body twitch/make nervous movements	*46:* Parts of my body twitch/make nervous movements
58: I pick my skin or other parts of my body	*63:* I would rather be with older people than people my own age
66: I repeat certain acts over and over	*66:* I repeat certain acts over and over
70: I see things that other people think aren't there	*70:* I see things that other people think aren't there
83: I store up things that I don't need	*36:* I accidentally get hurt a lot

84: I do things that other people think are strange	*84:* I do things that other people think are strange
85: I have thoughts that other people would think are strange	*85:* I have thoughts that other people would think are strange

Appendix 4: Word and non-word items (in French) used in the self-monitoring task (chapter 3)

Words	Non-words	Non-word source
c–v–c–v		
cage	leca	cale
cave	pypa	papy
date	daso	soda
dire	seva	vase
gare	cevi	vice
mode	becu	cube
note	seru	ruse
page	peju	jupe
vide	tolo	loto
lire	tevo	vote
rare	lijo	joli
lune	tevi	vite
c–v–c–v–c		
filer	fitel	filet
future	tacab	tabac
gamin	ciliv	civil
hiver	manir	marin
lapin	bunit	butin
laver	dinav	divan
jeter	binal	bilan
mener	vatel	valet
repas	satul	salut
repos	senol	selon
salon	terin	tirer
finir	furem	fumer
c–v–c–v–c–v		
cabine	migane	gamine
nature	lupile	pilule
navire	valage	lavage
balade	rimane	marine
pilote	luvome	volume
rapide	rapade	parade
refuge	nafile	finale
visite	cirane	racine
samedi	nabane	banane
valise	gacire	cigare
facile	necima	cinéma
copine	gadege	dégage

Appendix 5: Negative, neutral and positive word-items (in French) used in the reality-monitoring task (chapter 5)

Negative words	Neutral words	Positive words
détruire	timbre	douceur
cercueil	entretien	beauté
cancer	minute	liberté
crime	cravate	anniversaire
blessure	fauteuil	bonheur
danger	orange	bonté
drame	fenêtre	réussir
brutal	article	gentil
mourir	siècle	caresse
triste	modèle	rire

Appendix 6: Instructions and examples from The Movie for the Assessment of Social Cognition (MASC)

Note: For the purposes of reproducing the task material, verbal descriptions of the movie scenes are presented. Research subjects are presented with actual movie scenes and not a narrative describing the movie scene.

Instructions:
• You will be watching a 15 minute film. Please watch very carefully and try to understand what each character is feeling or thinking.
• Now you will meet each character: Sandra, Michael, Betty, and Cliff (a photo is shown of each)
• The film shows these four people getting together for a Saturday evening.
• The movie will be stopped at various points and some questions will be asked. All of the answers are multiple choice and require one option to be selected from a choice of four. If you are not exactly sure of the correct answer, please guess.
• When you answer, try to imagine what the characters are feeling or thinking at the very moment the film is stopped.
• The first scene is about to start. Are you ready? Again, please watch very carefully because each scene will be presented only once.

Question 1:
Imagine a movie scene which starts with the doorbell ringing. A young and attractive woman named Sandra opens the front door. Upon opening the door, a man, who looks to be around the same age as Sandra, enters the house. Sandra says "Hi" and the man asks her whether she is surprised. Before she can answer, he tells her that she looks terrific. He asks whether she did something with her hair. Sandra touches her hair and starts to say something but the young man compliments her by telling her that her hair looks very classy. The movie then stops and the following question is presented with four options to choose from:
What is Sandra feeling?
(1) that her hair does not look nice (no mentalizing)
(2) that she is pleased about his compliment (hypomentalizing)

(3) that she is exasperated about the man coming on too strong (hypermentalizing)

(4) that she is flattered but somewhat taken by surprise (accurate mentalizing)

Question 5:

In a previous scene, Sandra is on the phone with her good friend Betty, whom she implores to join them for dinner. Betty had previously stated that she could think of better things to do on a Saturday night and the scene ended. This scene starts with Sandra saying to Betty while smiling "Betty, I swear if you are not at this dinner on Saturday night, I will never ever speak to you again." The movie then stops and the following question is presented with four options to choose from:

Why is Sandra saying this?

(1) if Betty will not come, she will not speak to her anymore (hypomentalizing)

(2) to try to blackmail Betty into coming on Saturday (hypermentalizing)

(3) to persuade Betty in a joking way to come (accurate mentalizing)

(4) because Betty has better things to do on Saturday (no mentalizing)

Question 30:

All four characters are now in the kitchen preparing dinner together. The scene begins with Cliff asking Sandra for a bottle opener for the new bottle of wine. Michael then states that he has finished cutting all the onions and asks what else goes into the sauce that they are preparing. Betty checks with Sandra "two cups of cream, right?" and Michael looks over to Betty and responds: "If it were up to you you'd go for five, right?" The scene ends with Betty's sigh and expression of displeasure. The movie then stops and the following question is presented with four options to choose from:

What is Betty feeling?

(1) hates Michael and wants him to leave (hypermentalizing)

(2) five cups of cream would be too much for the sauce (no mentalizing)

(3) offended by Michael's comment (accurate mentalizing)

(4) astonished that Michael knows she likes cream (hypomentalizing)

Appendix 7: Examples of video frames of social interactions in The Movie for the Assessment of Social Cognition (MASC)

Cliff is the first one to arrive at Sandra's house for the dinner party. He and Sandra seem to enjoy themselves when Cliff is telling about his vacation in Sweden.

Question: Why is Sandra asking this?

When Michael arrives, he dominates the conversation, directing his speech to Sandra alone.

Slightly annoyed by Michael's bragging story, Sandra looks briefly in Cliff's direction and then asks Michael: "Tell me, have you ever been to Sweden?"

a) to integrate Cliff into the conversation *(correct)*

b) to get back to the topic of Sweden *(undermentalising: ,no ToM')*

c) to find out if Michael was in Sweden too *(undermentalising: ,reduced ToM')*

d) to be able to compare the two men *(overmentalising)*