Review

Attachment measures in middle childhood and adolescence: A systematic review of measurement properties

Tom Jewell[a,b,*], Tessa Gardner[c], Karima Susi[d,e], Kate Watchorn[f], Emily Coopey[d], Mima Simic[b], Peter Fonagy[f], Ivan Eisler[b]

[a] Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College, London, United Kingdom
[b] Child and Adolescent Eating Disorders Service, South London and Maudsley NHS Foundation Trust, London, United Kingdom
[c] Anna Freud Centre, London, United Kingdom
[d] Leicestershire Partnership NHS Trust, CAMHS Eating Disorders Team, Leicester, United Kingdom
[e] University of Oxford, Oxford Institute of Clinical Psychology Training, Oxford, United Kingdom
[f] University College London, Research Department of Clinical, Educational and Health Psychology, London, United Kingdom

HIGHLIGHTS

- Attachment insecurity is known to be a risk factor and correlate of psychopathology.
- This review is the first to systematically appraise the psychometric properties of all attachment measures in middle childhood and adolescence.
- Only a small number of attachment measures developed for middle childhood and adolescence have adequate measurement properties.
- Interview and projective measures of attachment should be used cautiously in clinical practice.
- Our findings converge with evidence from both infancy and adulthood in pointing to the underlying structure of attachment as being dimensional.

ARTICLE INFO

ABSTRACT

Background: Attachment theory proposes that humans develop representations of self and other in early childhood which are relatively stable across the life-course, and play a key role in psychological adaptation. However, to date, the psychometric properties of attachment measures in middle childhood and adolescence have not been evaluated in a systematic review.
Method: A systematic review (PROSPERO ID: CRD42017057772) was conducted using COSMIN criteria. Two researchers independently searched MEDLINE, PsychINFO and Embase databases for relevant articles.
Results: Fifty-four studies were included in the review. The methodological quality of studies was typically fair or poor, with only a small number of studies being rated as of good or excellent quality. The measurement properties of attachment measures in this age group were frequently rated as inadequate according to COSMIN criteria. The Child Attachment Interview (CAI) has the best psychometric properties of the interview and projective measures, and the Inventory of Parent and Peer Attachment (IPPA) the best evidence of the self-report measures. Overall, the evidence for the CAI and IPPA included both positive and negative findings relating to

adequacy of measurement properties.

Conclusions: Attachment measures in middle childhood and adolescence currently have limited evidence for the adequacy of their psychometric properties.

1. Attachment measures in middle childhood and adolescence: a systematic review of measurement properties

Attachment theory constitutes an evolving body of work developed by dozens of researchers and theoreticians over approximately sixty years. Two propositions are central to the theory: (1) patterns of attachment developed in infancy are relatively stable across the lifespan; (2) attachment patterns can help to explain the development of psychopathology (Bowlby, 1973, 1980). Both propositions can be tested empirically, but this requires the presence of valid and reliable attachment measures. Evidence points to attachment as a risk factor for the development of a range of psychopathology, including aggression and externalising behaviour (Fearon, Bakermans-Kranenburg, van IJzendoorn, Lapsley, & Roisman, 2010), internalizing disorders (Groh, Roisman, van IJzendoorn, Bakermans-Kranenburg, & Fearon, 2012), eating pathology (Jewell et al., 2016; Caglar-Nazali et al., 2014), and suicidality (Fergusson, Woodward, & Horwood, 2000). However, the power of attachment to predict later psychopathology is generally weak, raising doubts about its centrality as a causal factor (Fonagy, Luyten, & Allison, 2015). Moreover, there have been concerns that measures of attachment in childhood may have high levels of error (Fearon & Roisman, 2017), thus hampering efforts to understand the role of attachment in psychopathology.

Middle childhood and adolescence constitute developmental phases for which no 'gold standard' measures exist (Bosmans & Kerns, 2015). The aim of this study is to provide the first ever systematic review of the psychometric properties of all measures of attachment in middle childhood and adolescence. To understand the theoretical underpinnings of such measures, it is necessary to provide an account of the development of measurement approaches to infant and adult attachment. We note that the psychometric properties of these instruments have yet to be subjected to a systematic review, but recent narrative reviews of infant and adult attachment measures have been provided by Solomon and George (2016) and Crowell, Fraley, and Roisman (2016) respectively.

1.1. Measuring attachment in infancy and adulthood

The empirical assessment of human attachment began with Ainsworth, Blehar, Waters, and Wall's (1978) 'strange situation procedure' (SSP), in which trained observers rated infant behaviour in response to separation from and reunion with their primary caregiver. Although infants were rated across various scales, a discriminant function analysis suggested two underlying attachment dimensions, named *avoidance* and *anxiety*. Noting the clustering of the data in three groups resulted in the naming of three attachment categories: A, B and C, which came to be known as insecure-avoidant, secure and insecure-ambivalent. Later, Main and Solomon (1986) identified a fourth category in the SSP, disorganized attachment (D). The resulting four-category 'ABCD' paradigm directly informed the development of the Adult Attachment Interview (AAI) (George, Kaplan, & Main, 1985). This was originally developed in an attempt to explain the SSP classifications of infants by identifying differences in mother's representations of their own childhood experiences of being cared for. The four AAI categories - secure, dismissing, preoccupied and disorganized - were conceived of as explaining secure, avoidant, resistant and disorganized attachment patterns (respectively) in infants.

In a separate development, researchers such as Hazan and Shaver (1987) developed self-report measures of romantic *attachment style*. These measures were based on putative parallels between adults' ways of relating in romantic relationships and the concept of infant attachment categories. Brennan, Clark, and Shaver (1998) undertook a factor analysis of a large pool of items taken from self-report attachment style measures, and identified two underlying factors, *avoidance* and *anxiety*, echoing the dimensions identified by Ainsworth et al. (1978). In recent years, taxometric analyses of both the SSP (Fraley & Spiker, 2003) and the AAI (Fraley & Roisman, 2014) have also suggested that attachment is distributed across these two dimensions, rather than falling within categories. Nevertheless, concordance between AAI categories and self-reported attachment is trivial to low (Roisman et al., 2007).

Thus, the two main approaches to measuring adult attachment may be tapping related but distinct constructs, although the precise boundaries of the adult attachment construct are hard to operationalize (Allen, Stein, Fonagy, Fultz, & Target, 2005). The concept of adult attachment style as tapped by self-report-measures has been defined as a constellation of knowledge, expectations and insecurities that people hold about themselves and their close relationships (Fraley & Roisman, 2018). By contrast, the AAI seems to access internal working models of childhood caregiving experiences (Stein, Jacobs, Ferguson, Allen, & Fonagy, 1998), although Allen and Miga (2010) have suggested that the AAI may be best conceived as a measure of emotion regulation in the context of discussions about caregiving experiences.

1.2. Measuring attachment in middle childhood and adolescence

As can be gleaned from this overview, whilst measures of attachment exist for both infancy and adulthood, there are enormous differences not only in measurement approach, but also in the latent construct that is being assessed. In infancy, an observational measure is used to assess behaviour in young children who are developmentally reliant on their caregivers for survival, and are at a very early stage in their emotional, cognitive and social development. By contrast, attachment measures in adulthood assess mental representations, expressed through language, in adults who have acquired formal operational thinking (Inhelder & Piaget, 1958), and whose close relationships serve quite different functions, potentially including reproduction, child-rearing and emotional support. Approaches to assessing attachment in middle childhood and adolescence therefore must take into account the 'moving target' of child development, both in terms of the evolving function of the attachment system, and the changing abilities of the child.

Researchers have tried to access children's attachment representations using three main methods. Firstly, various self-report measures have been developed, including specific measures developed for middle childhood (Kerns et al., 1996), measures developed with both adults and adolescents in mind (Feeney, Noller, & Hanrahan, 1994), and downward extensions of adult romantic attachment measures (e.g. Brenning, Soenens, Braet & Bosmans, 2011). Secondly, interview approaches informed by the AAI have been developed, such as the CAI (Shmueli-Goetz, Target, Fonagy & Datta, 2008). Finally, some researchers have developed *projective* measures, in which children's stories and play in response to attachment-related prompts (e.g. Green, Stanley, Smith, & Goldwyn, 2000; Cassidy, 1988) are rated by trained coders. Thus far, attempts to critically appraise the measurement properties of these various measures have been limited. The National Institute for Health and Care Excellence (NICE) (2015) conducted the most thorough review so far, including measures of attachment from infancy to the age of eighteen, but excluding self-report measures. This review recommended the CAI for middle childhood, and the AAI for adolescents aged over fifteen. Kerns, Schlegelmilch, Morgan, and

Table 1
Characteristics, strengths and weaknesses of observer-rated measures.

Measure (Abbreviation)	# Questions	# Scales (range of scores)	Admin. time	Final categories or dimensions yielded	Focus of measure	Strengths	Weaknesses
Adult Attachment Interview (AAI)	20 questions	17 subscales: 5 experience scales (e.g. loving/ unloving); 3 states of mind regarding parents (e.g. idealising); 9 for state of mind with respect to attachment (e.g. coherence) (1–9)	1 h	Categories: Secure, Dismissing, Preoccupied, Unresolved/Disorganized, Cannot Classify	Adult attachment status	Well-validated in adults thus good choice for older adolescents.	Lengthy administration and coding time; known to be hard to achieve reliability for coding.
Adult Attachment Projective (AAP)	8 line drawings; 1 warm-up and 7 attachment scenes	Agency of self, connectedness, synchrony, deactivation, cognitive disconnection, segregated systems (?)	?	Categories: Secure, Dismissing, Preoccupied, Unresolved	Adult attachment status	Good inter-rater reliability.	Training for reliability is lengthy (9 days) thereby limiting its use in clinical settings.
Adolescent Attachment Prototype Questionnaire (AAPQ)	4 attachment prototype descriptions rated by clinician	Secure, Dismissing, Preoccupied, Incoherent/ Disorganized (1–5 Likert scale)	?	Dimensional and categorical ratings: Secure, Dismissing, Preoccupied, Incoherent/Disorganized	Generalized attachment status	Straightforward clinician-rated measure; no need to transcribe.	Inadequate construct validity (hypothesis testing).
Attachment Questionnaire – Adolescent Version (AQ-A)	37 clinician-rated items	Secure, Dismissing, Preoccupied, Incoherent/ Disorganized (1–7 Likert scale)	?	Dimensional and categorical ratings: Secure, Dismissing, Preoccupied, Incoherent/Disorganized	Generalized attachment status	Straightforward clinician-rated measure; no need to transcribe.	No data on construct validity (hypothesis testing).
Attachment Story Completion Task (ASCT)	Warm-up, then 5 story stems	Secure, Avoidant, Ambivalent, Disorganized (1–5)	?	Categories: Secure, Avoidant, Ambivalent, Disorganized	Parent attachment	Suitable for middle childhood.	Inadequate inter-rater reliability and construct validity (hypothesis testing; content and scoring methods have differed between studies.
Attachment Script Assessment (ASA)	4 stories (2 mother, 2 father)	Secure base script knowledge scale (1–7)	?	Continuous secure base script knowledge score.	Parent attachment	Good inter-rater reliability; simpler scoring method compared with other observer-rated measures.	Requires further study to explore convergent validity with other established measures; test retest reliability is unknown.
Bird's Nest Drawing (BND)	Task: Drawing of a bird's nest	Vitality, General pathology, Bizarreness, Optimism	?	Categories: Secure, Insecure	Parent attachment	Good inter-rater reliability; non-verbal means of assessment.	Inadequate construct validity; no evidence available for most psychometric properties.
Can You Tell Me? (CYTM)	22 questions	N/A	40 mins	10 categories based on DMM attachment strategies	Parent attachment	Semi-structured interview format can elicit rich clinical data	Inadequate psychometric properties; lengthy administration and coding
Child Attachment and Play Assessment (CAPA)	10 questions, 8 stems	N/A	?	6 categories based on DMM attachment strategies	Parent attachment	Semi-structured interview format can elicit rich clinical data.	Inadequate psychometric properties; lengthy administration and coding.
Child Attachment Interview (CAI)	15 questions	Involving anger, Idealization, Dismissal, Resolution of conflict, Balance of positive and negative references to attachment figures, Overall coherence, Emotional openness, Use of examples (1–9)	20–80 min	Categories: Secure, Dismissing, Preoccupied, Disorganized	Mother and father attachment	Overall the best-validated observer-rated measure of attachment in middle childhood and adolescence; semi-structured interview format can elicit rich clinical data.	Lengthy administration and coding time; known to be hard to achieve reliability for coding; inter-rater reliability is inadequate in most studies.
Computerized MacArthur Story Stem Battery (CMSSB)	6 story stems	Avoidance, Coherence, intentionality (1 − 12)	?	3 dimensions: Avoidance, Coherence, intentionality	Parent attachment	Computerized format could allow for easier administration to larger research samples.	Limited evidence of adequate psychometric properties.
Computerized Manchester Child Attachment Story Task (CMCAST)	4 story stems	N/A	?	Categories: Secure, Dismissing, Preoccupied, Disorganized	Parent attachment	Computerized format could allow for easier administration to larger research samples.	Inadequate inter-rater reliability.
Family Attachment Interview (FAI)	?	Secure, Fearful, Preoccupied, Dismissing (1–9)	1–2 h	Categories: Secure, Fearful, Preoccupied, Dismissing	Generalized attachment style	Semi-structured interview format can elicit rich clinical data.	Inadequate psychometric properties; lengthy administration and coding.
Friends and Family Interview (FFI)	27 questions	Coherence, Mentalization, Secure base availability, Social competence, School competence, Self-esteem, Peer relations, Sibling relations, Anxieties and defence, Differentiations of parental representations, Attachment classifications, Non-verbal codes (1–4)	?	Categories: Secure-autonomous, Insecure-dismissing, Insecure-preoccupied, Disoriented-disorganized	Parents, siblings and peers	Semi-structured interview format can elicit rich clinical data.	Inadequate psychometric properties; lengthy administration and coding.

Table 1 (*continued*)

Measure (Abbreviation)	# Questions	# Scales (range of scores)	Admin. time	Final categories or dimensions yielded	Focus of measure	Strengths	Weaknesses
Goal-Corrected Partnership Adolescent Coding System (GPACS)	Task: two 5-min conflict discussions (1 topic chosen by adolescent, 1 topic chosen by caregiver)	Collaborative communication, Warmth/ valuing, Caregiver validation of adolescent's voice, Caregiver's hostile/punitive/devaluing behaviour, Caregiver's role confused behaviour, Caregivers' odd or out-of-context/ contradictory behaviour, Caregiver's distracted, disoriented, or inwardly absorbed behaviour, Adolescent advances own view, Adolescent's hostile, punitive, or devaluing behaviour, Adolescent's caregiving, organizing, or entertaining behaviour, Adolescent's odd, out-of-context, or contradictory behaviour, Adolescent's distracted, disoriented, or inwardly absorbed behavior	10 mins	Dimensions: Security, Role confused, Hostile, Disorganized	Observed parent-adolescent interaction	Observational measure with dyadic focus; brief administration time; adequate inter-rater reliability.	Inadequate construct validity.
Goal-Corrected Partnership Adolescent Coding System (GPACS)	Task: two 5-min conflict discussions (1 topic chosen by adolescent, 1 topic chosen by caregiver)	Collaborative communication, Warmth/ valuing, Caregiver validation of adolescent's voice, Caregiver's hostile/punitive/devaluing behaviour, Caregiver's role confused behaviour, Caregivers' odd or out-of-context/ contradictory behaviour, Caregiver's distracted, disoriented, or inwardly absorbed behaviour, Adolescent advances own view, Adolescent's hostile, punitive, or devaluing behaviour, Adolescent's caregiving, organizing, or entertaining behaviour, Adolescent's odd, out-of-context, or contradictory behaviour, Adolescent's distracted, disoriented, or inwardly absorbed behavior	10 mins	Dimensions: Security, Role confused, Hostile, Disorganized	Observed parent-adolescent interaction	Observational measure with dyadic focus; brief administration time; adequate inter-rater reliability.	Inadequate construct validity.
Manchester Child Attachment Story Task (MCAST)	6 vignettes (1 warm-up, 5 attachment vignettes)	33 scales (most on 9-point scales)	20–30 min	Categories: Secure, Dismissing, Preoccupied, Disorganized, Cannot Classify	Parent attachment	Suitable for younger end of middle childhood.	Psychometric properties generally inadequate.
School-age Assessment of Attachment (SAA)	7 picture cards	N/A	?	13 categories based on DMM attachment strategies	Parent attachment	Semi-structured interview format can elicit rich clinical data.	Inadequate psychometric properties; lengthy administration and coding.
Secure Base Script Test (SBST)	5 sets of word prompts (1 warm-up, 2 parents, 2 best friends)	Scriptedness (1–7)	8–20 min	1 dimension: 7-point Scriptedness score	Generalized attachment script	Relatively brief administration; good inter-rater reliability.	Requires further study to explore convergent validity with other established measures; test retest reliability is unknown.
Separation Anxiety Test (SAT)	9 photographs	Attachment (5–20) Self-reliance (4–16) Avoidance (9–27)	?	3 dimensions: attachment, self-reliance, avoidance	Generalized attachment	Suitable for younger end of middle childhood.	Psychometric properties generally inadequate.

Table 2
Characteristics, strengths and weaknesses of self-report measures.

Measure (Abbreviation)	# items	Scales	Response options	Range of scores	Focus of measure	Strengths	Weaknesses
Adolescent Attachment Questionnaire (AAQ)	9	Angry distress / Availability / Goal-corrected partnership	1–5	?	Parents	Brief measure, adequate structural validity.	Inadequate internal consistency; doubtful content validity.
Adolescent Friendship Attachment Scale (AFAS)	30	Secure / Anxious / Avoidant	1–5	1–5	Peers	Adequate internal consistency, test retest reliability, content validity.	Inadequate structural validity.
Adolescent Friendship Attachment Scale - Short Form (AFAS-SF)	15	Secure / Anxious / Avoidant	1–5	1–5	Peers	Brief, adequate internal consistency.	Inadequate structural validity.
Attachment Style Questionnaire (ASQ)	40	Confidence; Relationships as secondary; Discomfort with closeness; Preoccupation with relationships; Need for approval;	1–6	40–240	Generalized	Adequate internal consistency; applicable to both adolescents and adults; widely used in adult studies.	Inadequate structural validity.
Adolescent Unresolved Attachment Questionnaire (AUAQ)	9	Failed protection / Anger/dysregulation / Fear	1–5	?	Parents	Adequate content and structural validity.	Inadequate internal consistency.
Experiences in Close Relationships - Revised for Children (ECR-RC)	36	Avoidance / Anxiety	1–7	?	Parent	Adequate internal consistency and hypothesis testing; based on widely used adult attachment style measure.	Inadequate structural validity.
Experiences in Close Relationships – Relationship Structures (ECR-RS)	9	Avoidance / Anxiety	1–7	6–42 (Avoidance) / 3–21 (Anxiety)	Parent, best friend, romantic partner	Adequate internal consistency; based on widely used adult attachment style measure.	Inadequate hypothesis testing; uncertain structural validity.
Experiences in Close Relationships – Revised – General Short Form (ECR-R-GSF)	20	Avoidance / Anxiety	1–5	1–5	Generalized	Adequate content validity and internal consistency; based on widely used adult attachment style measure.	Inadequate hypothesis testing and structural validity.
Inventory of Parent and Peer Attachment (IPPA)	81 (28# each parent, 25# peers)	Communication / Trust / Alienation	1–5	16–92 (Mother, Father) / 19–82 (Peer)	Mother, father, peers	Studied relatively extensively across a range of samples and countries; adequate internal consistency and hypothesis testing.	Lengthy; inadequate structural validity.
Inventory of Parent and Peer Attachment- Brief version (IPPA-B)	28 (14# mother, 14# father)	Communication / Trust / Alienation	1–5	?	Parent	Brief version of a widely-studied measure; adequate internal consistency, structural validity and hypothesis testing.	Only validated in one study of fair quality; only available in French.
Inventory of Parent and Peer Attachment – 45 (IPPA-45)	45 (15# each for mother, father, peers)	Communication / Trust / Alienation	1–5	?	Mother, father, peers	Adequate internal consistency; shorter version of a widely-studied measure.	Inadequate structural validity.
Inventory of Parent and Peer Attachment – Revised (IPPA-R)	81 (28# each parent, 25# peers)	Communication / Trust / Alienation	1–5	16–92 (Mother, Father) / 19–82 (Peer)	Mother, father, peers	Adequate hypothesis-testing and content validity.	Lengthy; inadequate internal consistency and structural validity.
Inventory of Parent and Peer Attachment – Short Form (IPPA-S)	24 (12# parent, 12# peer)	Communication / Trust / Alienation	1–5	?	Parent, peers	Brief version of a widely-studied measure.	Inadequate internal consistency, structural validity and hypothesis testing.
Preoccupied and Avoidant Coping Questionnaire (PACQ)	36	Avoidant / Preoccupied	0–2	0–36 (Avoidant) / 0–32 (Preoccupied)	Parent	Adequate internal consistency.	Inadequate content validity and hypothesis testing.
People In My Life (PIML)	44 (20# parent, 24# peer)	Communication / Trust / Alienation	1–5	?	Parent, peers	Based on widely-studied IPPA; adequate hypothesis testing and structural validity.	Inadequate internal consistency.
Security Scale (SS)	15	Security	1–4	1–4	Parent	Brief measure developed for middle childhood.	Conflicting findings about adequacy of internal consistency and hypothesis testing.

Abraham (2005) and Wilson and Wilkinson (2012) have also conducted narrative reviews but neither included ratings of the methodological quality of studies, nor did they make recommendations about which measures have the best psychometric properties.

1.3. Rationale for this review

Attachment theory constitutes a highly important paradigm within the field of child development and psychopathology, resulting in thousands of empirical papers. However, the reliability and validity of measures is fundamental to the conduct and interpretation of this body of research.

1.4. Objective

Our primary aim in undertaking this review is to make recommendations about which attachment measures have the best psychometric properties, thereby providing a guide to researchers and clinicians in the field. We will also identify gaps in the evidence and make recommendations about promising avenues for future research.

2. Method

2.1. Literature search

This systematic review was registered with PROSPERO (CRD 42017057772) and completed in accordance with PRISMA guidelines (Moher, Liberati, Tetzlaff & Altman, 2009). Two independent researchers searched MEDLINE, PsychINFO and Embase databases for relevant articles up to the end of June 2017. Eligibility criteria were: (1) English language; (2) published in a peer-reviewed journal; (3) aim of study is to develop a measure, or evaluate the properties of a measure, that assesses attachment in children or adolescents using either interview, task, projective method or self-report; (4) participants in the study are aged between 6 and 18 years inclusive; (5) measure is theoretically derived from attachment theory. The search strategy was constructed by TG, and refined by TJ, over three waves of literature searching in February 2016, November 2016 and June 2017. The

search strategy is publicly available at the PROSPERO protocol registration for this review.

2.2. Data extraction

This study used the COSMIN checklist (De Vet, Terwee, Mokkink, & Knol, 2011; materials available at https://www.cosmin.nl), a tool for systematic reviews of measurement properties. Data were extracted using templates from the COSMIN tool. Characteristics of studies (e.g. sample size, age of participants) can be found in Appendices A and B for observer-rated measures and self-report measures respectively. Characteristics of the measures themselves (e.g. number of items and scales) can be found in Tables 1 and 2.

2.3. Assessment of measurement properties

Pairs of reviewers independently assessed each study using the COSMIN checklist for the following criteria: internal consistency, test-retest and inter-rater reliability, content validity, hypotheses testing (i.e. construct validity), cross cultural validity and criterion validity. Measurement error and responsiveness to change are part of the COSMIN checklist but were not included in this review since no studies were found addressing these measurement properties.

Each study was rated by a pair of reviewers. TJ rated all papers, and TG, KW, KS and EC rated approximately a quarter of papers each. Disagreements were resolved by PF. Each measurement property was rated on a four-point rating scale (poor, fair, good, excellent) for methodological quality of the study, thereby assessing risk of bias within studies. Each measurement property also received a rating for the adequacy of the measurement property (assigned as '+', '−' or '?'). COSMIN criteria for the adequacy of measurement properties can be found in Table 3. Ratings of the methodological quality and adequacy of measurement properties within individual studies are in Appendices C and D (observer-rated and self-report measures respectively).

2.4. Data synthesis

A synthesis of the strength of evidence for each measurement

Table 3
COSMIN Criteria for adequacy of measurement properties.

Content validity (including face validity)	+	All items refer to relevant aspects of the construct to be measured AND are relevant for the target population AND are relevant for the purpose of the measurement instrument AND together comprehensively reflect the construct to be measured
	?	Not all information for '+' reported
	−	Criteria for '+' not met
Structural validity	+	Unidimensionality: EFA: First factor accounts for at least 20% of the variability AND ratio of the variance explained by the first to the second factor > 4 OR
		Bi-factor model: Standardized loadings on a common factor > 0.30 AND correlation between individual scores under a bi-factor and unidimensional model > 0.90
		Structural validity: CFI or TLI or comparable measure > 0.95 AND (Root Mean Square Error of Approximation (RMSEA) < 0.06 OR Standardized Root Mean Residuals (SRMR) < 0.08)
	?	Not all information for '+' reported
	−	Criteria for '+' not met
Internal consistency	+	At least limited evidence for unidimensionality or positive structural validity AND Cronbach's alpha(s) ≥ 0.70 and ≤ 0.95
	?	Not all information for '+' reported OR conflicting evidence for unidimensionality or structural validity OR evidence for lack of unidimensionality or negative structural validity
	−	Criteria for '+' not met
Reliability (Internal and Test-Retest)	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	−	Criteria for '+' not met
Construct validity (hypothesis testing)	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with instrument(s) measuring related construct(s) AND no differences between relevant groups reported
	−	Criteria for '+' not met
Cross-cultural validity	+	No important differences found between language versions in multiple group factor analysis or DIF analysis
	?	Multiple group factor analysis AND DIF analysis not performed
	−	One or more criteria for '+' not met
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70
	?	Not all information for '+' reported
	−	Criteria for '+' not met

property was conducted for all measures other than those which only achieved ratings of poor methodological quality (Tables 4 and 5: observer-rated and self-report measures respectively). Criteria used to define the strength of evidence can be found in Appendix E. Finally, we incorporated a brief narrative summary of strengths and weaknesses of each measure into Tables 1 and 2, so that researchers and clinicians can rapidly appraise the characteristics of different measures 'at a glance'.

3. Results

Our search yielded 601 articles once duplicates were removed (see Fig. 1). TG and KW separately screened by title and abstract, then assessed eligibility in 101 full-text articles. Disagreements at screening stage were resolved by TJ. Reference lists were checked for additional articles not picked up in the search stage. Fifty-four relevant articles were identified (see Appendix F for references of included studies).

3.1. Internal consistency

Adequate internal consistency (alpha > 0.7) was reported in studies of the following self-report measures: AFAS, AFAS-SF, ASQ, ECR-RC, ECR-RS, ECR-R-GSF, IPPA-B and IPPA-45. There were studies reporting both adequate and inadequate internal consistency for the SS, PACQ and IPPA. Study quality was generally fair, but was excellent in the case of the AFAS, AFAS-SF, ASQ, ECR-RS and IPPA-45.

For observer-rated measures, several studies examined internal consistency across questions and sub-scales (ASA, ASCT, CAI, FFI, SAT, SBST). Only the AQ-A and ASA reported adequate internal consistency, both in studies of poor methodological quality for the evaluation of this property.

3.2. Test-retest reliability

Test-retest reliability data was notably lacking for most measures. It was reported for the AAQ, AFAS, AUAQ, PACQ, CAI and MCAST but no studies met COSMIN adequacy criteria (ICC or kappa > 0.7).

3.3. Inter-rater reliability (interview, observation and projective measures)

Adequate inter-rater reliability was reported for the AAI, AAP, BND, CAI, GPACS and two measures using the secure base script paradigm, the ASA and SBST. The CAI demonstrated good inter-rater reliability in the study by Borelli et al. (2016), but not in earlier studies.

3.4. Content validity

Content validity proved hard to assess using the standard COSMIN

criteria, thus these were adapted. Studies were rated as excellent for methodological quality if (1) they demonstrated evidence of iterations in the development of the measure, (2) if the assumptions underlying the measure were tested (e.g. through piloting), and (3) if they involved an expert panel. Face validity of the measure in terms of its theoretical links to attachment theory were also considered. Positive ratings of content validity were given to the AFAS, CAI, ECR-R-GSF and IPPA in studies of good or excellent methodological quality, and also to the AUAQ, CMCAST, MCAST, SAA and SAT in studies of fair or poor methodological quality.

3.5. Structural validity

Several studies evaluated factor structure using exploratory and confirmatory factor analysis. Adequate structural validity was reported for the AAQ, AUAQ, IPPA-B and PIML.

For observer-rated measures, factor analyses were conducted, typically using subscales in the place of items. In most such studies the results were given an indeterminate rating ('?') since the findings were not interpretable within the COSMIN scheme. These indeterminate ratings were given to structural validity studies of the ASCT, GPACS, MCAST and CAI, with the exception of the Zachrisson, Røysamb, Oppedal, and Hauser (2011) study of the CAI, which was given a positive rating for adequate measurement properties and rated as excellent for methodological quality. This study is notable in that the confirmatory factor analysis demonstrated adequate model fit, as defined by COSMIN, for a two factor model comprised of avoidance and preoccupation. One other notable study of structural validity was that by Waters et al. (2015). This was a taxometric study pointing to the dimensional structure of attachment as measured by the Attachment Script Assessment.

3.6. Hypotheses testing

Under the COSMIN scheme, various aspects of construct validity, such as convergent and discriminant validity, are assessed under the banner of 'hypotheses testing'. More favourable ratings of methodological quality are assigned for studies that test multiple, specific hypotheses including the direction and magnitude of correlations. Of the observer-rated measures, ratings of adequate construct validity were assigned to studies of the AAI, AAP and CMSSB. The ASA, CAI, SAA and SAT showed inconsistent findings, with both positive and negative studies of hypotheses testing reported. For self-report measures, adequate ratings were given to studies of the AFAS-SF, ECR-RC, IPPA, IPPA-B, IPPA-R and PIML. Findings for the SS were inconsistent.

Table 4

Synthesis of psychometric properties and level of evidence for observer-rated measures.

Measure	Internal consistency	Test retest reliability	Inter-rater reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Criterion validity
AAI			?			+		
AAP			+ +			+		
AAPQ						−		
ASCT	?		− − −		?	− −		
ASA	?		+ +		+ +	+/−		
BND			+ +			−		
CAI	?	− −	+/−	+ + +	+ + +	+/−		
CMSSB			?			+		
CMCAST			−	+				−
GPACS			+ +		?	− −		
MCAST	−	?	− −	+	?	−	?	
SBST	?		+ +			− −		
SAT	?	?	?	?		+/−		

+ + + or − − −, strong evidence; + + or − −, moderate evidence; + or −, limited evidence; ?, unknown; blank cell, no evidence available.

Table 5

Synthesis of psychometric properties and level of evidence for self-report measures.

Name of measure	Internal consistency	Test retest reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Criterion validity
AAQ	− − −	−	−	+ + +	−	?	
AFAS	+ + +	?	+ +	− − −	− −		
AFAS-SF	+			−	+		+
AUAQ	− − −	?	+	+ + +	−	?	
ASQ	+ + +			− − −			
ECR-RC	+			−	+		
ECR-RS	+ + +			?	−		
ECR-R-GSF	+		+ +	−	−		
IPPA	+/−		+ +	−	+ +	?	
IPPA-45	+ + +			− − −			
IPPA-R	− −			−	+ +		
IPPA-S	−			−	−		
IPPA-B	+			+	+	?	
PIML	−			+	+		
PACQ	?	?	−		− −		
SS	?				+/−		

+ + + or − − −, strong evidence; + + or − −, moderate evidence; + or −, limited evidence; +/−, conflicting findings; ?, unknown; blank cell, no evidence available.
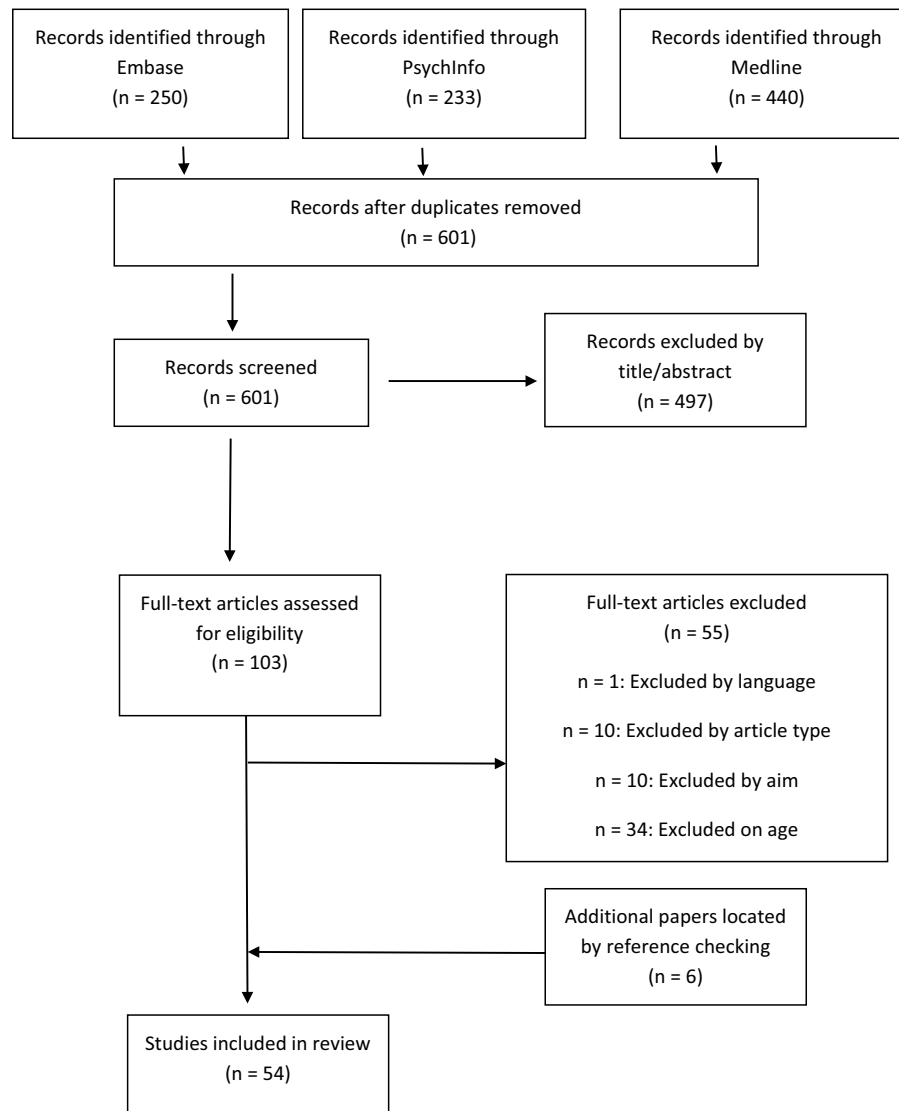

Fig. 1. Flowchart for search strategy.

3.7. Cross-cultural validity

Few studies specifically investigated cross-cultural validity as defined within the COSMIN taxonomy, which refers to the degree to which the performance of a translated or culturally-adapted instrument is an adequate reflection of its performance in its original version. Where it was investigated, study quality was rated as poor. Adequate cross-cultural validity was demonstrated for the FFI. An indeterminate rating was given in studies of the AAQ, AUAQ, IPPA, IPPA-B, and MCAST.

3.8. Criterion validity

Studies of criterion validity were rare, which is unsurprising in a field lacking an accepted 'gold standard' measure. The AFAS-SF demonstrated adequate criterion validity against the AFAS, whilst the CMCAST did not do so against the MCAST.

3.9. Synthesis of results

Our synthesis of the strength of the evidence for psychometric properties is presented in Tables 4 and 5 (observer-rated and self-report, respectively). Overall, no measure has demonstrated consistent evidence of good psychometric properties across a range of criteria. However, our findings point to the CAI and IPPA currently having the best evidence of adequate measurement properties. The CAI has positive findings in support of its content validity, structural validity when assessed using two dimensions (Zachrisson et al., 2011), and various positive findings relating to construct validity (e.g. Borelli et al., 2016). However, its inter-rater reliability is sub-optimal in most studies, with the exception of that by Borelli et al. (2016). The IPPA exists in several versions, none of which have emerged as demonstrating adequacy across a range of psychometric properties. In general, the structural validity of the measure is inadequate, and does not accord with the two-dimensional structure that our review suggests is most strongly supported by evidence. Further, its findings on internal consistency have been mixed. However, it has demonstrated adequate construct validity across a relatively large number of studies compared to other self-report measures.

4. Discussion

Overall our review points to a lack of evidence of adequate measurement properties for most available attachment measures in middle childhood. However, we wish to draw attention to some important points that should be borne in mind when interpreting our findings.

Firstly, the COSMIN tool yields categories, with necessarily arbitrary values chosen as cut-offs to distinguish adequate from inadequate measurement properties. In some cases the statistical values that led to a negative rating were close to the value required for a positive rating. This point applies similarly to the ratings of methodological quality, in which COSMIN operates a 'worst score counts' algorithm. This means that the final rating of methodological quality is defined by the lowest score obtained for that measurement property; thus a single flaw could lead to a rating of 'fair' when it would otherwise have been rated 'excellent'. We applied a similar rule in rating the adequacy of a measurement property where data for several subscales were presented: one sub-optimal value was enough to lead to a negative rating of adequacy for that property.

The implication of all of these points is that a quick glance at our ratings may lead to an underestimation of both the adequacy of measurement properties and also the methodological quality of the evidence. Finally, readers should note that we followed the COSMIN guidance around study selection, including only studies that specifically stated the investigation of measurement properties as an aim of the study (De Vet et al., 2011). Previous reviews (e.g. Kerns & Brumariu,

2016) have included a broader range of studies as providing evidence of validity, such as studies looking at associations between attachment and emotional regulation (e.g. Brumariu, Kerns, & Seibert, 2012). However, including studies that did not specifically aim to examine psychometric properties would have increased the risk of bias, led to an unwieldy number of studies for review, and made it harder for future researchers to reproduce our review. Nevertheless, we accept that by using COSMIN we have taken a relatively stringent approach to the selection and rating of studies.

We believe that this review has helpfully summarised the state of evidence on psychometric properties for individual measures. In addition, this review also allows us to consider some broader questions that we consider to be fundamental to the measurement and conceptualisation of attachment in middle childhood and adolescence. We have organised our discussion around two key questions, before moving on to recommendations for research and clinical practice.

4.1. What is the most valid and reliable approach to assessing attachment in middle childhood and adolescence?

In general we do not consider, a priori, that any measurement approach is inherently more appropriate to the measurement of attachment across middle childhood and adolescence. One exception to this is the assessment of attachment in the early phase of middle childhood (age 6–8 years) in which self-report measures are unlikely to be valid as a consequence of children's more limited reading and cognitive ability. Unfortunately, our review points to the relatively poor psychometric properties of measures developed for this age group (e.g. Green et al., 2000). The CAI was initially developed for children aged 7–13 years, and currently has the best evidence of psychometric properties for early middle childhood. We believe a study of psychometric properties of the CAI in a sample of 6–9 year olds would be worthwhile, as currently there is a lack of evidence for measures demonstrating adequate psychometric properties for this age group.

In older middle childhood and adolescence, interviews constitute a well-validated measurement approach. Whilst the CAI has been studied in adolescence (Venta, Shmueli-Goetz & Sharp, 2014) clinicians and researchers may want to consider whether the AAI or AAP might be more appropriate for older teenagers. Only one measure in our review, the GPACS (Obsuth, Hennighausen, Brumariu, & Lyons-Ruth, 2014), utilised an observation of adolescent-caregiver interaction, with scales measuring dyadic interaction, in addition to adolescent and caregiver behaviour within a ten-minute task discussing areas of disagreement. In addition to capturing the quality of in-vivo parent-caregiver interaction, the measure is theoretically-informed by a conceptualisation of disorganized attachment in adolescence. As such, the GPACS assesses more extreme features of disorganized attachment (such as role reversal) that are less likely to be captured by other measures.

Also promising are measures of secure base scriptedness (Dykas, Woodhouse, Cassidy, & Waters, 2006; Psouni & Apetroaia, 2014), in which the task involves creating stories using word prompts, which have demonstrated relatively strong evidence of adequate psychometric properties. Such measures have numerous advantages including shorter administration time and simpler scoring method when compared to interview measures such as the CAI. Further studies examining convergent validity of scriptedness measures with the best-performing interview and self-report measures would be an important contribution to the field.

Finally, the reliability, and especially validity, of self-report measures of attachment is important to consider, not least since NICE (2015) did not include such measures in their guideline on attachment. Most self-report measures in this review did not examine convergent validity with interview or projective assessment methods. Two exceptions are the IPPA and SS, which have both been found to be correlated with attachment as measured by the CAI (Borelli et al., 2016). Importantly, however, the correlations were below the 0.4 cut-off which is

used by convention as evidence that two measures are tapping the same construct. Thus, strictly speaking, we can conclude that some self-report measures of attachment are correlated with attachment interviews, but we cannot be certain they are rating the same construct. For adolescents, measures such as the ECR and ASQ can be used to assess attachment styles, which can be assumed to have conceptual continuity with the attachment style construct as measured in adults. Such measures are needed in order to shed light on the developmental antecedents of adult attachment styles, for instance through longitudinal studies (e.g. Jones et al., 2018). However, as we discuss later, there is scope for the improvement of self-report measures.

4.2. Is attachment distributed categorically or continuously?

The four-category ABCD paradigm has held a central place in attachment theory and measurement approaches. In our review, interview and projective measures based on this paradigm typically reported sub-optimal structural validity and inter-rater reliability (kappa < 0.7). By contrast, dimensional approaches to scoring such measures demonstrated favourable reliability (e.g. Psouni & Apetroaia, 2014; Waters et al., 2015). Importantly, Zachrisson et al. (2011) found evidence of a factor structure comprising two dimensions underlying the CAI. Thus the findings from this review appear to converge with emerging findings at other points in the lifespan, with taxometric analyses of both the Strange Situation Procedure (Fraley & Spieker, 2003) and the Adult Attachment Interview (Fraley & Roisman, 2014) supporting the idea of attachment being distributed across two dimensions, rather than four categories.

If attachment is distributed continuously, are self-report measures yielding continuous scales the best approach to measuring it? Unfortunately, almost all self-report measures in this review demonstrated sub-optimal structural validity. This review included various self-report measures, often based on adult attachment style measures. These measures have often been subject to numerous revisions, such as changes to wording and item length (e.g. ECR and IPPA studies). Despite exhaustive factor analysis using large samples, these measures have failed to meet criteria for good structural validity as defined by COSMIN. This raises difficult questions for the field. Does the lack of structural validity reflect problems with the measures themselves, or are the constructs they set out to measure not reflective of the phenomenology of attachment in middle childhood and adolescence? Likewise, does attachment in this age group not break down into the ABCD categories, or are available measures not able to detect them reliably? Based on the evidence in this review, it seems plausible that the attachment construct in middle childhood and adolescence is inherently difficult to measure reliably. This may be because attachment representations themselves are relatively fluid at this age (Jones et al., 2018); perhaps also because the developing nature of children's cognitive and socio-emotional abilities presents challenges in capturing such a complex construct.

Thus, the underlying structure of attachment in middle childhood and adolescence is unclear, with neither the ABCD model for interview/ projective measures, nor the two factor (avoidance/anxiety) structure of adult attachment style measures demonstrating strong evidence of validity in this age group. This has important implications for research and clinical practice.

4.3. Implications for research

In keeping with the findings by NICE (2015), our study highlights the relatively poor methodological quality of many studies in the field. However, within our review we note with encouragement a trend towards improved study quality over time. Some key methodological principles worth highlighting are: clearly stated hypotheses that include predictions about the direction and magnitude of expected correlations, and reporting on both the amount of missing data and how it was handled. Studies of test-retest reliability and sensitivity to change are

required, as are studies investigating attachment measures in a range of different sociocultural contexts. Recent studies conducted in Africa and Asia (Sochos & Lokshum, 2017; Wan, Danquah, & Mahama, 2017) are a welcome development for the field; more such studies are needed.

For interview measures based on the ABCD paradigm (e.g. CAI, MCAST), research on simpler coding systems yielding dimensional scores of avoidance and preoccupation, would lead to a number of benefits. These include improved inter-rater reliability, theoretical congruence with other developments in the field (Fraley and Roisman, 2014; Fearon and Roisman, 2017), and increased statistical power in research, such as longitudinal designs investigating the impact of attachment on developmental outcomes (e.g. Wright, Hill, Sharp, & Pickles, 2018).

Our review casts doubt on the notion that there is a single latent attachment construct which is tapped by all the measures in this review, given both the heterogeneity in measurement approaches, and the evidence surveyed on convergent validity. Like Bosmans and Kerns (2015), we agree that it is more fruitful to ask what aspect of attachment one is trying to assess, rather than what the 'gold standard' attachment measure might be. 'Attachment and affiliation' are already included in the Research Domains Criteria (Cuthbert, 2014), and in our view this presents an opportunity to advance the measurement of attachment across the lifespan. Researchers should aim to develop developmentally appropriate measures of more precise, well-validated lower-order attachment constructs, (e.g. secure base scriptedness) which ultimately belong to higher-order domains relating to socio-emotional processes and reward (Fonagy & Luyten, 2018). Rather than privileging attachment above other constructs, it would be more helpful to place attachment within a broader project to improve the science of developmental psychopathology and mental health treatment. Understanding how different aspects of attachment, during different developmental phases, play a role within broader social processes, has greater potential to lead to innovations in treatment than working with poorly-validated concepts and measures.

In order to develop a more empirically-supported approach to attachment in middle childhood and adolescence, structural equation modelling (SEM) could be used to investigate the extent to which different measures, and indeed individual items, load on to latent attachment variables. Ideally this would be undertaken with large datasets in which a variety of measurement approaches have been taken, including measures of attachment and associated constructs such as mentalizing and emotion regulation. Exploratory work of this kind should be possible with existing data sets, although the ideal would be to design studies with large samples that could simultaneously investigate the structure of attachment and evaluate new measures. This should include the development of new self-report measures, the starting point for which must be a clear conceptualisation of what is meant by attachment. In studies in which attachment has been measured by a variety of methods, the process of refining the questionnaire should include the extent to which items load on to a latent attachment variable (or variables) in SEM. By this means, the measure would avoid the pitfalls of shared method variance and reliance on adult attachment style models for concurrent validity.

Our review also highlights a clear need for research on measurement error. Fearon and Roisman (2017) recommend the development of better attachment measures that allow direct assessment of the relationship between indicators, error and underlying constructs. Once again, studies employing structural equation modelling could investigate the degree of measurement error for each item in a self-report questionnaire, or each scale in an interview/projective measure, by assessing correlations with a latent attachment variable.

4.4. Implications for clinical practice

Attachment theory is an influential theoretical framework in many clinical contexts such as psychological therapies with children and families. In the UK, national guidelines published by NICE (2015) specifically advise clinicians working with children in child protection and

adoption settings to consider attachment in their assessment and treatment planning, and recommend the MCAST, CAI and AAI. Our findings suggest that such measures are vulnerable to high measurement error, as evidenced by unfavourable ratings of inter-rater reliability in many studies. We therefore suggest that, when used, the findings of such measures should be interpreted tentatively as a clinical hypothesis, and understood as being informative only as one aspect of a much broader assessment of a child and their caregivers; moreover, this hypothesis may need to be reviewed over time as new information emerges. In court settings, evidence provided from such measures that has been rated by only one clinician should not be seen as authoritative, since a second rater may well disagree on the assignment of attachment category. Finally, we encourage clinicians, policy-makers and members of the public not to reify attachment categories. Studies in this review suggest that attachment status is not necessarily predictive of psychopathology in children and adolescents. For instance, in studies using the CAI, rates of secure attachment in clinical samples have been as high as 30% in adolescent psychiatric inpatients (Venta et al., 2014), whilst Scott, Riskman, Woolgar, Humayun, and O'Connor (2011) reported security in 52% and 73% respectively in moderate and high-risk samples for conduct problems. This compares with rates of approximately 60% attachment security in normative child samples (Shmueli-Goetz et al., 2008; Green et al., 2000).

5. Limitations

As discussed earlier, our review applied stringent inclusion criteria. As a consequence, our review may under-report the breadth of evidence on psychometric properties for attachment measures in middle childhood and adolescence, although excluded studies are likely to be of low-quality. Our review excluded studies where the mean age of participants lay outside of the 6–18 age range. Consequently, our review excluded some measures that are suitable at the extremities of this age range, such as the Attachment Style Interview (Schimmenti & Bifulco, 2015), which has been validated in a sample of 16–25 year-olds. Furthermore, we were not able to appraise the risk of bias arising from publication bias across studies. Studies in our review reported a range of different statistics across a variety of measurement properties; as such, there was no valid and reliable way to assess publication bias. Given the likely researcher bias towards publishing positive results, it is possible that our review over-estimates the adequacy of psychometric properties across measures, since there may be unpublished data showing negative findings.

6. Conclusion

The field of attachment is entering an exciting phase in which new empirical and theoretical insights are emerging. Longitudinal studies of attachment stability suggest that attachment may be less stable in adolescence than in adulthood (Jones et al., 2018), and that genetic influences on attachment may come in to play in adolescence that were not present in infancy (Fearon et al., 2014). Recent theories have suggested that the early stages of sexual maturation (andrenarche) may constitute a 'switch-point' in the development of attachment strategies (Del Giudice, Angeleri, & Manera, 2009), and that adult attachment styles may be more influenced by recent interpersonal experiences than distal, early caregiving experiences (Fraley & Roisman, 2018). The attachment field is reliant on the availability of valid, reliable and sensitive measures that can be used to test theories and build evidence. In the clinical realm, good measures are needed to test both aetiological models in which childhood attachment experiences are implicated (e.g. Fonagy, Gergely, Jurist, & Target, 2002) and also the role of attachment in treatments (e.g. Diamond et al., 2010). In adult clinical samples, attachment is a predictor of psychotherapy outcome (Levy, Kivity, Johnson, & Gooch, 2018) and is associated with differential response to treatment across a range of disorders including psychosis (Carr, Hardy, & Fornells-Ambrojo, 2018; Gumley, Taylor, Schwannauer, & MacBeth, 2014) and eating disorders (Tasca and Balfour,

2014). The role of attachment in the process of psychological treatments for children and adolescents is under-researched, and the further development of psychometrically sound measures will help to advance understanding in this area.

Selecting a suitable attachment measure, whether in a clinical or research context, is a complex matter. Our review provides various important sources of information that can guide the decision including measurement approach (i.e. interview, task, self-report), administration time, and the type of attachment relationship that is assessed. In particular, we advise close scrutiny of face validity: that is, clinicians and researchers need to assess the extent to which the conceptualisation of attachment and how it is assessed fit with the purposes for which the measure is being chosen.

In summary, our review suggests that there are currently large gaps in our knowledge of the psychometric properties of attachment measures, with the lack of data on sensitivity to change being particularly regrettable. We found limited evidence of adequate psychometric properties, but identified the CAI and IPPA as currently having the best evidence of such properties amongst observer-rated and self-report measures respectively. The ASA, a measure of secure base scriptedness, was identified as a promising measure worthy of future research on psychometric properties. Our findings point to the advantages of dimensional rather than categorical approaches to measurement, with more favourable inter-rater reliability and structural validity ratings observed in measures yielding dimensional scores. Future studies are needed that test specific hypotheses and that shed light on the underlying structure of attachment representations in middle childhood and adolescence.

Contributions

TJ, IE and PF wrote the protocol for the systematic review. TJ, TG and KW conducted literature searches. TJ, TG, KS, EC and KW rated the included studies. TJ wrote the first draft of the manuscript and all authors contributed to and have approved the final manuscript.

Conflict of interest

All authors report no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2018.12.004.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Allen, J. G., Stein, H., Fonagy, P., Fultz, J., & Target, M. (2005). Rethinking adult attachment: A study of expert consensus. *Bulletin of the Menninger Clinic, 69*(1), 59–80.

Allen, J. P., & Miga, E. M. (2010). Attachment in adolescence: A move to the level of emotion regulation. *Journal of Social and Personal Relationships, 27*(2), 181–190.

Borelli, J. L., Somers, J., West, J. L., Coffey, J. K., Los Reyes, A., & Shmueli-Goetz, Y. (2016). Associations between attachment narratives and self-report measures of attachment in middle childhood: Extending evidence for the validity of the Child Attachment Interview. *Journal of Child and Family Studies, 25*, 1235–1246.

Bosmans, G., & Kerns, K. A. (2015). Attachment in middle childhood: Progress and prospects. *Newdirections for Child and Adolescent Development, 2015*(148), 1–14.

Bowlby, J. (1973). *Attachment and loss, vol. 2: Separation: Anxiety and anger.* New York, NY: Basic Books.

Bowlby, J. (1980). *Attachment and loss: Vol. 3: Loss.* New York, NY: Basic Books.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult

attachment: An integrative overview. In J. A. Simpson, & W. S. Rholes (Eds.). *Attachment theory and close relationships* (pp. 46–76). New York: The Guilford Press.

Brenning, K., Soenens, B., Braet, C., & Bosmans, G. (2011). An adaptation of the experiences in close relationships scale-revised for use with children and adolescents. *Journal of Social and Personal Relationships, 28*, 1048–1072.

Brumariu, L. E., Kerns, K. A., & Seibert, A. (2012). Mother–child attachment, emotion regulation, and anxiety symptoms in middle childhood. *Personal Relationships, 19*(3), 569–585.

Caglar-Nazali, H. P., Corfield, F., Cardi, V., Ambwani, S., Leppanen, J., Olabintan, O., & Micali, N. (2014). A systematic review and meta-analysis of 'Systems for Social Processes' in eating disorders. *Neuroscience & Biobehavioral Reviews, 42*, 55–92.

Carr, S. C., Hardy, A., & Fornells-Ambrojo, M. (2018). Relationship between attachment style and symptom severity across the psychosis spectrum: A meta-analysis. *Clinical Psychology Review, 59*, 145–158.

Cassidy, J. (1988). Child-mother attachment and the self in six-year-olds. *Child development*(pp. 121–134).

Crowell, J. A., Fraley, R. C., & Roisman, G. I. (2016). Measurement of individual differences in adult attachment. In J. Cassidy, & P. R. Shaver (Eds.). *Handbook of attachment: Theory, research, and clinical applications* (pp. 598–635). (3rd ed.). New York, NY: Guilford Press.

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*, 28–35.

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine.* Cambridge: Cambridge University Press.

Del Giudice, M., Angeleri, R., & Manera, V. (2009). The juvenile transition: A developmental switch point in human life history. *Developmental Review, 29*, 1–31.

Diamond, G. S., Wintersteen, M. B., Brown, G. K., Diamond, G. M., Gallop, R., Shelef, K., & Levy, S. (2010). Attachment-based family therapy for adolescents with suicidal ideation: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*, 122–131.

Dykas, M. J., Woodhouse, S. S., Cassidy, J., & Waters, H. S. (2006). Narrative assessment of attachment representations: Links between secure base scripts and adolescent attachment. *Attachment & Human Development, 8*(3), 221–240.

Fearon, R. P., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Lapsley, A. M., & Roisman, G. I. (2010). The significance of insecure attachment and disorganization in the development of children's externalizing behavior: A meta-analytic study. *Child Development, 81*, 435–456.

Fearon, P., Shmueli-Goetz, Y., Viding, E., Fonagy, P., & Plomin, R. (2014). Genetic and environmental influences on adolescent attachment. *Journal of Child Psychology and Psychiatry, 55*(9), 1033–1041.

Fearon, R. P., & Roisman, G. I. (2017). Attachment theory: Progress and future directions. *Current Opinion in Psychology, 15*, 131–136.

Feeney, J. A., Noller, P., & Hanrahan, M. (1994). Assessing adult attachment. In M. B. Sperling, & W. H. Berman (Eds.). *Attachment in adults: Clinical and developmental perspectives* (pp. 128–152). New York: Guilford Press.

Fergusson, D. M., Woodward, L. J., & Horwood, L. J. (2000). Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood. *Psychological Medicine, 30*, 23–39.

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self.* New York, NY: Other Press.

Fonagy, P., & Luyten, P. (2018). Conduct problems in youth and the RDoC approach: A developmental, evolutionary-based view. *Clinical Psychology Review, 64*, 57–76.

Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders, 29*, 575–609.

Fraley, R. C., & Roisman, G. I. (2014). III. Categories or dimensions? A taxometric analysis of the adult attachment interview. *Monographs of the Society for Research in Child Development, 79*, 36–50.

Fraley, R. C., & Roisman, G. I. (2018). The development of adult attachment styles: Four lessons. *Current Opinion in Psychology.* https://doi.org/10.1016/j.copsyc.2018.02.008 Available in early view.

Fraley, R. C., & Spieker, S. J. (2003). Are infant attachment patterns continuously or categorically distributed? A taxometric analysis of strange situation behavior. *Developmental Psychology, 39*, 387.

George, C., Kaplan, N., & Main, M. (1985). *The adult attachment interview. Unpublished protocol.* Berkeley: Department of Psychology, University of California.

Green, J., Stanley, C., Smith, V., & Goldwyn, R. (2000). A new method of evaluating attachment representations in young school-age children: The Manchester Child Attachment Story Task. *Attachment & Human Development, 2*, 48–70.

Groh, A. M., Roisman, G. I., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Fearon, R. P. (2012). The significance of insecure and disorganized attachment for children's internalizing symptoms: A meta-analytic study. *Child Development, 83*, 591–610.

Gumley, A. I., Taylor, H. E. F., Schwannauer, M., & MacBeth, A. (2014). A systematic review of attachment and psychosis: Measurement, construct validity and outcomes. *Acta Psychiatrica Scandinavica, 129*(4), 257–274.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology, 52*, 511.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jewell, T., Collyer, H., Gardner, T., Tchanturia, K., Simic, M., Fonagy, P., & Eisler, I.

(2016). Attachment and mentalization and their association with child and adolescent eating pathology: A systematic review. *International Journal of Eating Disorders, 49*, 354–373.

Jones, J. D., Fraley, R. C., Ehrlich, K. B., Stern, J. A., Lejuez, C. W., Shaver, P. R., & Cassidy, J. (2018). Stability of attachment style in adolescence: An empirical test of alternative developmental processes. *Child Development, 89*, 871–880.

Kerns, K. A., & Brumariu, L. E. (2016). Attachment in middle childhood. In J. Cassidy, & P. R. Shaver (Eds.). *Handbook of attachment: Theory, research, and clinical applications* (pp. 349–365). New York, NY: Guilford Press.

Kerns, K. A., Klepac, L., & Cole, A. (1996). Peer relationships and preadolescents' perceptions of security in the child-mother relationship. *Developmental Psychology, 32*, 457–466.

Kerns, K. A., Schlegelmilch, A., Morgan, T. A., & Abraham, M. M. (2005). Assessing attachment in middle childhood. In K. A. Kerns, & R. A. Richardson (Eds.). *Attachment in middle childhood* (pp. 46–70). New York: Guilford Press.

Levy, K. N., Kivity, Y., Johnson, B. N., & Gooch, C. V. (2018). Adult attachment as a predictor and moderator of psychotherapy outcome: A meta-analysis. *Journal of Clinical Psychology, 74*(11), 1996–2013.

Main, M., & Solomon, J. (1986). Discovery of an insecure-disorganized/disoriented attachment pattern. In T. B. Brazelton, & M. W. Yogman (Eds.). *Affective development in infancy* (pp. 95–124). Westport, CT: Ablex.

National Institute for Health and Care Excellence (2015). *Children's attachment: Attachment in children and young people who are adopted from care, in care or at high risk of going into care. NICE guideline 26.* London: The British Psychological Society & The Royal College of Psychiatrists.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*(4), 264–269.

Obsuth, I., Hennighausen, K., Brumariu, L. E., & Lyons-Ruth, K. (2014). Disorganized behavior in adolescent-parent interaction: Relations to attachment state of mind, partner abuse, and psychopathology. *Child Development, 85*, 370–387.

Psouni, E., & Apetroaia, A. (2014). Measuring scripted attachment-related knowledge in middle childhood: the Secure Base Script Test. *Attachment & human development, 16*(1), 22–41.

Roisman, G. I., Holland, A., Fortuna, K., Fraley, R. C., Clausell, E., & Clarke, A. (2007). The adult attachment interview and self-reports of attachment style: An empirical rapprochement. *Journal of Personality and Social Psychology, 92*, 678.

Schimmenti, A., & Bifulco, A. (2015). Linking lack of care in childhood to anxiety disorders in emerging adulthood: The role of attachment styles. *Child and Adolescent Mental Health, 20*(1), 41–48.

Scott, S., Riskman, J., Woolgar, M., Humayun, S., & O'Connor, T. G. (2011). Attachment in adolescence: Overlap with parenting and unique prediction of behavioural adjustment. *The Journal of Child Psychology and Psychiatry, 52*, 1052–1062.

Shmueli-Goetz, Y., Target, M., Fonagy, P., & Datta, A. (2008). The Child Attachment Interview: A psychometric study of reliability and discriminant validity. *Developmental Psychology, 44*, 939.

Sochos, A., & Lokshum, C. (2017). Adolescent attachment in Nepal: Testing the factorial validity of two scales. *European Journal of Developmental Psychology, 14*(4), 498–508.

Solomon, J., & George, C. (2016). The measurement of attachment security and related constructs in infancy and early childhood. In J. Cassidy, & P. R. Shaver (Eds.). *Handbook of attachment: Theory, research, and clinical applications* (pp. 366–396). (3rd ed.). New York, NY: Guilford Press.

Stein, H., Jacobs, N. J., Ferguson, K. S., Allen, J. G., & Fonagy, P. (1998). What do adult attachment scales measure? *Bulletin of the Menninger Clinic, 62*(1), 33.

Tasca, G. A., & Balfour, L. (2014). Attachment and eating disorders: A review of current research. *International Journal of Eating Disorders, 47*(7), 710–717.

Venta, A., Shmueli-Goetz, Y., & Sharp, C. (2014). Assessing attachment in adolescence: A psychometric study of the child attachment interview. *Psychological Assessment, 26*(1), 238–255.

Wan, M. W., Danquah, A. N., & Mahama, S. (2017). Child-caregiver Attachment Representations in a Non-Western Context: The Feasibility and Cultural Equivalence of Story Stems in Urban Ghana. *Social Development, 26*, 59–79.

Waters, T. E., Fraley, R. C., Groh, A. M., Steele, R. D., Vaughn, B. E., Bost, K. K., ... Roisman, G. I. (2015). The latent structure of secure base script knowledge. *Developmental Psychology, 51*, 823.

Wilson, J. M., & Wilkinson, R. B. (2012). The self-report assessment of adolescent attachment: A systematic review and critique. *Journal of Relationships Research, 3*, 81–94.

Wright, N., Hill, J., Sharp, H., & Pickles, A. (2018). Maternal sensitivity to distress, attachment and the development of callous-unemotional traits in young children. *Journal of Child Psychology and Psychiatry.* https://doi.org/10.1111/jcpp.12867 Available on Early View.

Zachrisson, H. D., Røysamb, E., Oppedal, B., & Hauser, S. T. (2011). Factor structure of the child attachment interview. *European Journal of Developmental Psychology, 8*, 744–759.

Tom Jewell, the first author, is a Clinical Doctoral Research Fellow funded by the National Institute of Health Research (NIHR) in the United Kingdom. His doctoral study is investigating attachment and mentalization as predictors of outcome in family therapy for adolescent anorexia nervosa. He works clinically as a family therapist in adolescent eating disorders.