More Haste Less Speed: A Meta-Analysis of Thinking Latencies During Planning in People with Psychosis

Andrew J. Watson*, Eileen M. Joyce*, Andrew J. B. Fugardb, Verity C. Leesonc, Thomas R. E. Barnesc and Vyv Huddyb*

Running head: Thinking latencies in people with psychosis

aSobell Department of Motor Neuroscience and Movement Disorders, University College London, London, WC1N 3BG, UK. Andrew.J.Watson@ucl.ac.uk

bDepartment of Clinical, Educational and Health Psychology, University College London. London, WC1E 6BT, UK. V.huddy@ucl.ac.uk

cCentre for Psychiatry, Imperial College London, London W12 0NN, UK

* Vyv.huddy@ucl.ac.uk, Department of Clinical, Educational and Health Psychology, University College London. London, WC1E 6BT, UK.
Abstract

Cognitive impairment is a core feature of psychosis, with slowed processing speed thought to be a prominent impairment in schizophrenia and first-episode psychosis. However, findings from the Stockings of Cambridge (SOC) planning task suggest changes in processing speed associated with the illness may include faster responses in early stages of planning, though findings are inconsistent. This review uses meta-analytic methods to assess thinking times in psychosis across the available literature. Studies were identified by searching PubMed, Web of Science and Google Scholar.

Eligibility criteria: 1) included a sample of people with non-affective psychosis according to DSM III, DSM IV, DSM V or ICD-10 criteria; 2) employed the SOC task; 3) included a healthy control group; and 4) published in English. We identified 11 studies that employed the SOC task. Results show that people with psychosis have significantly faster initial thinking times than non-clinical participants, but significantly slower subsequent thinking times during problem execution. These findings indicate that differences in processing speed are not limited to slower responses in people with psychosis but may reflect a preference for step-by-step processing rather than planning before task execution. We suggest this style of responding is adopted to compensate for working memory impairment.

Key words: Schizophrenia; Cognition; Executive Function; Processing Speed, CANTAB
1. Introduction

People with psychosis show impaired cognitive performance at the time of the first episode of illness (Mesholam-Gately et al., 2009) and after multiple episodes (Dickinson et al., 2007). Compared to healthy controls, the level of impairment is substantial in almost all cognitive domains (Dickinson et al., 2007). This generalised pattern of impairments has been interpreted as reflecting a core impairment of schizophrenia (Dickinson and Harvey, 2009). One of these cognitive domains is processing speed, which can be defined as “the speed with which an individual can perform any cognitive operation” (Salthouse, 1996) and is usually measured as the number of correct responses achieved on a task within a given time. Evidence for slowed information processing has been consistently observed in those with a diagnosis of schizophrenia (Knowles et al., 2010; Nuechterlein, 1977) and non-affective first-episode psychosis (Mesholam-Gately et al., 2009; Mohamed et al., 1999). A prominent quantitative synthesis of the literature concluded that processing speed was the most impaired of all cognitive domains in schizophrenia (Dickinson et al., 2007). Impaired processing speed in schizophrenia is suggested as one of the “crucial mechanisms of impaired cognitive functioning” (Brebion et al., 2009), and is associated with illness risk (Reichenberg et al., 2010), and clinical (Leeson et al., 2010) and functional outcomes (Brekke et al., 1997; Gold et al., 2002).

Speed of information processing is widely assessed using basic measures such as the Digit Symbol Substitution Test (DSST) and the Trail Making Test (TMT), both of which contribute to the speed of processing domain of the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) battery (Nuechterlein et al., 2008). Morrens et al., (2007) suggest that, whilst these tests are sensitive to psychomotor slowing, they are also sensitive to a wide range of higher level cognitive functions, such as working memory or cognitive flexibility, with deficits in subsets of these functions potentially causing poor performance in these tasks. Indeed, faster response times in people with psychosis have been reported in planning tasks, although other studies have failed to find this. These findings contradict the suggestion that processing speed is central to the cognitive difficulties in people with psychosis, with patients often responding more quickly than healthy controls.
The aforementioned planning studies employed the computerised Stockings of Cambridge (SOC) planning task, a variation of the classic Tower of London problem (Shallice, 1982). In order to be successful, SOC requires participants to mentally plan their sequence of moves before beginning to complete them. Participants are provided with two different arrangements of 'balls' sitting in 'stockings' hanging from an imagined snooker or pool table; they are asked to plan and execute a series of moves on one arrangement to match the second displayed arrangement, according to a set of rules. This is known as the “plan and move” condition. Key to this task is that participants are asked to solve the problem in the minimum number of moves possible and not to begin until they know which moves to make. The problems vary in difficulty, reflecting the number of planned moves required to solve the problem accurately. The computerised nature of the task also allows a detailed assessment of performance latencies which provide a clue as to how individuals approach the task.

For example, there are 'yolked' motor control problems whereby the computer controls for individual motor ability by presenting participants with their own solutions to problems and then asking them to follow the exact same sequence of moves on the lower half of the screen (follow condition); by subtracting these 'motor' times from the 'planning' times, the amount of time a participant spends purely thinking about the task can be derived (discounting that slower responding is solely due to individual differences in motor function). Further, thinking times can be differentiated into ‘initial’ times (reflecting the length of time participants spend considering the problem solution before attempting it) and ‘subsequent’ times (reflecting the amount of time thinking about each subsequent move as they execute the solution). Initial thinking times are the difference in time between the participant selecting the first ball in the “plan and move” condition and selecting the first ball in the “follow” condition. Subsequent thinking times are calculated by taking the time between selection of the first ball and the completion of the task, and dividing it by the total number of moves made. This task provides a rigorous means of measuring processing speed impairments in people with psychosis versus healthy controls. The findings in the literature have been inconsistent, so a quantitative synthesis of the literature is warranted to determine if there is evidence of a combination of faster and slower thinking times during planning.
1.1 Aims of the Study

We carried out a systematic review and meta-analysis of the literature on the SOC task to 1) examine the overall impairment in planning accuracy and 2) establish if this is accompanied by group differences in initial and subsequent thinking times.

2. Method

2.1. Search Strategy

Studies were identified by searching PubMed, Web of Science and Google Scholar using the following search terms: (Cambridge Neuropsychological Test Automated Battery OR Stockings of Cambridge OR Tower of London OR Tower of Hanoi OR CANTAB OR TOL OR TOH OR SOC) AND (Psychosis OR Schizophrenia). We included the search terms of other planning tasks - Tower of London and Tower of Hanoi – to establish if the SOC task had been employed in any of these studies or if there was the possibility of mislabelling of the SOC task. This search was conducted for studies published until March 2016 and included congress abstracts.

2.2. Eligibility criteria

Studies were included if they 1) included a sample of people with schizophrenia or non-affective psychosis according to DSM III or DSM IV American Psychiatric Association (2000), DSM V American Psychiatric Association (2013) or ICD-10 (1992) criteria, 2) employed the CANTAB SOC task, 3) included a healthy (non-psychiatric) control group, and 4) were published in the English language. Two reviewers (VH and AW) independently screened and determined eligibility for included studies. Disagreements were resolved by discussion, with arbitration via third reviewer (EMJ) planned but not needed. To ensure the highest standard of reporting, we adopted “Preferred Reporting Items for Systematic Reviews and Meta-Analyses” (PRISMA) guidelines (Moher et al., 2009).

2.3. Data extraction and recorded variables
Two reviewers used standardised forms to independently extract data. We collected data on demographic variables reported in studies, including date of publication, sample size, age of participants and sex ratio. We also gathered data on the IQ of the psychosis and healthy control groups. Disagreements were dealt with as described above.

2.4 Risk of Bias

The CANTAB is a standardised computerised assessment tool, designed to minimise assessor bias. A remaining area of potential bias was inadequate matching of the two participant groups on demographic variables. For this reason, coded individual study variables that would enable the matching of clinical and healthy control groups to be assessed.

2.5 Calculating of standardised effect sizes

The SOC task has four conditions of problem complexity ranging from two to five moves required for perfect problem execution. There was inconsistency in how the variables were reported, with some studies reporting all four complexity levels, some fewer than four and with others reporting only an average – or composite - across conditions. We report the number of perfect solutions, the initial, and the subsequent thinking times for the lower difficulty level (3 move), higher difficulty level (5 move) and composite (2 – 5 move) conditions. These were the most commonly reported variables in the studies that were reviewed. Based on the data reported in the selected studies we estimated standardised effect size (SMD) as Hedges’ g (Hedges, 1981): the difference between the test performance (accuracy or response time) divided by the pooled standard deviation. The estimate for one study (Braw et al. 2008) revealed an SMD that was extremely large. We were unable to confirm with the authors if this was an error, so we used a ‘leave one out’ analysis (see below) that tests for undue influence of individual studies. A small number of effect sizes were obtained from statistics reported in studies following methods described by Thalheimer and Cook (2002). Better performance and longer thinking times are indicated by positive effect sizes.

2.5 Meta analytical procedure
We conducted 9 individual meta-analyses on the difference between people with psychosis and healthy controls on the following variables: number of perfect solutions, initial thinking time and subsequent thinking time. Random effects models were estimated using the metafor package (Viechtbauer, 2010) in R version 3.1.0 (R-Core-Team, 2014) (http://www.R-project.org/).

Heterogeneity of effects was estimated with the Q statistic (Hedge and Olkin, 1985) and \( I^2 \) (Higgins et al., 2003). We used guidance by Deeks, Higgins, and Altman (Deeks J, 2011) to determine the presence of substantial heterogeneity. Finally, we used funnel plots and trim-and-fill analyses to assess publication bias (Duval and Tweedie, 2000).

3. Results

3.1. Selection of articles

We found 387 studies, of which 11 met our criteria; these included 662 patients with psychosis and 497 healthy controls. Of the 387 reports, 292 were excluded because: 1) a non-affective psychosis sample was not included (\( n=149 \)); 2) the CANTAB/SOC task was not used (\( n=107 \)); 3) a case control design was not used (\( n=43 \)), the article was not in English or did not report data (\( n=25 \)) or a combination of these factors (see Figure 1). No studies using the DSM-V were identified. Five of the studies included participants with a diagnosis of schizophrenia only (Badcock et al., 2005; Braw et al., 2013; Kontis et al., 2013; Pantelis et al., 1997a; Tyson et al., 2004), three included a diagnosis of schizophrenia, schizophreniform or schizoaffective disorder (Hilti et al., 2010; Joyce et al., 2002; Leeson et al., 2009a), two included schizophrenia or other non-affective psychotic disorder (Braw et al., 2008; Fagerlund et al., 2006) and one specified “schizophrenia or non-organic and non-affective psychosis” (Saleem et al., 2013). Of the 11 eligible studies (see Table 1), two included some of the same participants (Braw et al., 2008; Braw et al., 2013) but the studies were separately analysed as different variables were reported: 5-move variables were reported in one of the studies while composite variables were reported in the other. Another of the eligible studies (Hilti et al., 2010) failed to report thinking latencies and included some data previously reported in a prior study. We obtained raw data from the authors so that non-overlapping effect sizes and thinking latencies could be reported.
3.2. SOC Performance (see Table 2)

There were significant differences between cases and controls at all difficulty levels. There was a very large effect of participant group at the 5-move level of difficulty (-1.61 (95% CI [-3.14, -0.08], \( p = 0.039 \)) and a moderate effect at both the 3-move level of difficulty (-0.58 [-0.75, -0.40], \( p < 0.001 \)) and the composite of all difficulty levels (-0.66 [-0.85, -0.46], \( p < 0.001 \)) (see Figure 2).

3.3 Analysis of initial thinking times (see Table 2)

The initial thinking time variables showed significantly shorter latencies in the psychosis groups at the 5-move problem level (-0.40 [-0.61, -0.20], \( p < 0.001 \)) (see Figure 3a) but not 3-move problems (0.22 [-0.09, 0.54], \( p = 0.186 \)). There were relatively fewer studies reporting 3-move versus 5-move data. The effect size of the difference for the composite initial thinking time was not statistically significant (\( p = 0.655 \)). There was significant heterogeneity at the 3-move level of difficulty but not the 5-move level.

3.4. Analysis of subsequent thinking times

For subsequent thinking times there were significantly longer latencies for 3, 5 and the composite variable in psychosis groups (see Figure 3b). There was no heterogeneity of effect sizes in either the 3-move, 5-move or composite problems.

3.5. Risk of bias: matching of healthy control groups

All studies employed healthy control groups that were matched for age and all but one matched for sex ratio. The majority of studies that reported IQ (4 out of 7 studies) employed healthy control groups which demonstrated significantly higher IQ than those in the psychosis groups. A moderation analysis was conducted for each of the nine outcomes to test the effect of whether groups were IQ matched. One of the nine outcomes was statistically significant (other \( p \)'s > 0.11), initial thinking times for 3 move problems [\( Q_{M}(1) = 7.7, p = 0.005 \)]. There was no difference between the psychosis group and control group for unmatched studies (\( k = 2, SMD = -0.08, 95\% CI [-0.27, 0.11], p = 0.41 \)). However, for matched studies, participants in the psychosis group were slower on initial
thinking than control group (k = 2, SMD = 0.53, 95% CI [0.14, 0.92], \(p = 0.007\)). For the other eight out of nine outcomes, there was no evidence of a differential effect of matching.

### 3.6 Sensitivity analyses

The participants with psychosis in one of the included studies (Hilti et al., 2010) were naïve to antipsychotic medication at the time of testing. We performed a leave-one-out analysis on all outcomes to test the impact on results. The pattern of results (direction of effect and whether the 95% CIs exclude zero) was identical for all but one analysis: the number of perfect solutions for 5 move problems \((k = 5)\). Removing Joyce (Joyce et al., 2002), Leeson (Leeson et al., 2009b), or Braw (Braw et al., 2008) rendered the \(p > 0.05\). However, this effect appears to be because of the Saleem data, noticeably outlying in the forest plot. Removing this study dramatically improves the precision of the estimate (SE = 0.08 without this study versus 0.78 when it is included). Furthermore, now the leave-one-out analysis for the remaining four studies had no impact on the pattern of results.

### 3.7 Publication bias

A trim and fill analysis was conducted to test for publication bias. The pattern of results (direction of effect and whether the 95% CIs exclude zero) was unaffected (see Figure 4). Seven of the nine effect sizes changed by less than 0.1. Of the other two, the largest was for initial thinking time on 3 move problems, and reduced the estimated effect size from 0.22 (95% CI [-0.09, 0.54], \(p = 0.17\)) to 0.04 (95% CI [-0.29, 0.38], \(p = .8\)). The second largest shift was for subsequent thinking time on 5-move problems where the effect size was reduced from 0.39 (95% CI [0.20, 0.57], \(p < 0.001\)) to 0.25 (95% CI [0.05, 0.46], \(p = 0.02\)). These data indicate very little evidence of publication bias.

### 4. Discussion

#### 4.1 Summary of evidence

Our meta-analysis confirmed that people with psychosis show abnormalities in planning with respect to both accuracy (i.e. number of perfect solutions) and thinking latencies. For the most difficult, 5-move problems, both initial and subsequent thinking times were significantly different in
patients compared to healthy controls: initial thinking times were significantly faster whilst subsequent thinking times were significantly slower. For the composite variables, initial thinking times were not different but subsequent thinking times remained slower in patients. These results were not influenced by noteworthy evidence of publication bias. The subsequent thinking time findings were consistent with the wider literature on slowing across a range of tasks. However, the deficit in subsequent thinking time was accompanied by faster initial response latencies for the most complex problems. This indicates that viewing the slowing of processing speed as a key feature of the cognitive profile of schizophrenia samples could be mistaken.

The current findings indicated that faster initial thinking time in patients was accompanied by slower subsequent thinking time. Thus, compared to healthy controls, those with psychosis showed a preference for step-by-step processing rather than first planning and then moving. The latter effect might be expected if an inadequately planned sequence of moves needed to be reordered into the correct sequence during execution, resulting in slower subsequent thinking time. The observation that controls made less errors than patients suggests that the longer initial thinking times ensures that the execution phase is focused on carrying out the moves that were imagined prior to beginning problem execution. In the one touch version of the SOC task, where execution involves only stating the number of required moves, people with schizophrenia show longer latencies (Huddy et al., 2007). The key difference with the current computerised version is that the task set-up allows the participant to progress towards a solution by trying out different possibilities by physically moving the balls on the screen. This activity provides a compensatory support to working memory that is not available in the one touch version. The changes in planning performance reported above in the corpus of studies, i.e. faster initial responses accompanied by increased errors, are inconsistent with a finding of equivalent reflection impulsivity in people with schizophrenia and healthy controls (Huddy et al., 2013). Whilst the current findings may appear to be indicative of impulsivity it is possible that abnormalities in planning reflect a compensatory strategy for poor working memory. Further research is required to disentangle these possibilities and to determine the role of working memory in the successful completion of the SOC task and how it relates to the measures of processing speed.
Faster initial thinking times in people with psychosis were not found across all levels of difficulty, as might be expected if there were global impulsivity. Instead, the initial thinking time differences were found only for the more difficult problem trials but not the easier 3-move problems. Consistent with this effect, two studies reported an interaction between problem difficulty and group so that controls took progressively more time to consider the solution before initiation, which was less evident in patients. This interaction can be understood as a failure to adequately increase thinking time as problems become more difficult in people with psychosis. The fact that the majority of studies missed this effect by reporting only isolated sub-test scores or global performance variables demonstrates how the full potential of the SOC task has not been realised by much of the research in this area.

4.2 Limitations

The majority of studies included in the review failed to match the healthy control group for pre-existing IQ differences leaving open the possibility that differences in intellectual ability could confound the results on speeded initial thinking times in 5-move problems in people with psychosis. However, there are several reasons to think that IQ differences do not substantially confound the results. First, the initial thinking time effect sizes for 5-move problems did not demonstrate significant heterogeneity across studies that employed matched or non-matched control groups. Secondly, sensitivity analysis using the leave one out procedure did not change our pattern of results. Furthermore, as noted in the introduction, the direction of the initial thinking time difference is in favour of faster thinking in people with psychosis suggesting that a single global impairment in cognitive processing, resulting in inaccuracy and slowed responses, is not a sufficient explanation for the pattern of findings reported here.

One inclusion criterion for the study was the employment of the SOC rather than any other measure of planning that also provided an estimate of thinking latencies. Thus, interpretation of our findings is limited to the SOC task as the measure employed; to assess generalisability future studies
should employ measures that index other forms of planning. However, the advantage of applying such a criterion is that it allows a clear interpretation of the meaning of the thinking time variable, as the tasks are identical in their computerised procedure so task administration differences are minimised. The validity and reliability of the measures could have been compromised by including studies where thinking times were gathered by hand. Another shortcoming of this review is that the majority of participants in the studies were prescribed medication at the time of testing, with one exception. However, the results were unchanged when this study was removed from the analysis.

4.3 Conclusions

In conclusion, the planning impairments found in people with psychosis compared with healthy controls are accompanied by both shorter initial and longer subsequent thinking times. This suggests that patients spend less time thinking before attempting the harder problems and take more time thinking before each subsequent move, but still make more errors. These data support cognitive remediation therapies that involve both education about cognitive processing changes that follow psychosis and training in strategies that overcome them. Faster initial thinking times in the context of impaired accuracy indicates a deficit in problem elaboration prior to execution of the task which may be subject to cognitive remediation. One ongoing clinical trial specifically targets processing speed using practice based protocol. However, the current findings suggests a strategy training approach is required as increased speed could be detrimental to performance. It is notable that cognitive remediation is effective for reducing impairments in processing speed in trials that use a strategy training approach. Strategy training targets improvements in the identification of core task variables, an explicit plan and execution the solution. This approach would necessarily entail slower, more often accurate, performance. Thus, performance on the SOC would be ideal for indexing change in cognitive remediation therapy.
References


<table>
<thead>
<tr>
<th>Publication</th>
<th>Year</th>
<th>N</th>
<th>Sex (% Male)</th>
<th>Age</th>
<th>IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pantelis, Barnes (Pantelis et al., 1997b)</td>
<td>1997</td>
<td>31</td>
<td>36</td>
<td>58.1</td>
<td>80.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>47.48</td>
<td>48.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>101.27</td>
<td>97.16</td>
</tr>
<tr>
<td>Joyce, Hutton (Joyce et al., 2002)</td>
<td>2002</td>
<td>81</td>
<td>136</td>
<td>60.5</td>
<td>78.7*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.1</td>
<td>25.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>104.64</td>
<td>99.67*</td>
</tr>
<tr>
<td>Tyson, Laws (Tyson et al., 2004)</td>
<td>2004</td>
<td>17</td>
<td>28</td>
<td>-</td>
<td>64.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>39.4</td>
<td>33.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>106.17</td>
<td>101.17</td>
</tr>
<tr>
<td>Badcock, Michie (Badcock et al., 2005)</td>
<td>2005</td>
<td>33</td>
<td>24</td>
<td>78.8</td>
<td>79.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34.7</td>
<td>32.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>108.3</td>
<td>101.42*</td>
</tr>
<tr>
<td>Fagerlund, Pasberg (Fagerlund et al., 2006)</td>
<td>2006</td>
<td>40</td>
<td>18</td>
<td>40</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.3</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110.8</td>
<td>87.9*</td>
</tr>
<tr>
<td>Braw, Bloch (Braw et al., 2008)</td>
<td>2008</td>
<td>44</td>
<td>44</td>
<td>61.4</td>
<td>77.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.6</td>
<td>24.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Leeson, Robbins (Leeson et al., 2009a)</td>
<td>2009</td>
<td>111</td>
<td>151</td>
<td>50.5</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.3</td>
<td>26.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>103.8</td>
<td>93.2*</td>
</tr>
<tr>
<td>Hilti, Delko (Hilti et al., 2010)</td>
<td>2010</td>
<td>33</td>
<td>26</td>
<td>72.7</td>
<td>82.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>Saleem, Harte (Saleem et al., 2013)</td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Kontis, Theochari (Kontis et al., 2013)</td>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td>78</td>
</tr>
<tr>
<td>Braw, Sitman (Braw et al., 2013)*</td>
<td>2013</td>
<td>37</td>
<td>101</td>
<td>83.8</td>
<td>72.3</td>
</tr>
</tbody>
</table>

* Indicates a significant difference between participants with psychosis (Psychosis) and healthy controls (HC).

*a These statistics refer to an overall group were collapsed across symptom subcategories reported in the paper.
Table 2. Summary of Meta Analyses

<table>
<thead>
<tr>
<th>Measure</th>
<th>Difficulty</th>
<th>Level</th>
<th>k</th>
<th>SMD</th>
<th>Lower</th>
<th>Upper</th>
<th>p</th>
<th>Q</th>
<th>p(Q)</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Thinking time</td>
<td></td>
<td>3</td>
<td>5</td>
<td>0.22</td>
<td>-0.09</td>
<td>0.54</td>
<td>0.168</td>
<td>11.2</td>
<td>0.025</td>
<td>68.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>7</td>
<td>-0.40</td>
<td>-0.61</td>
<td>-0.20</td>
<td>&lt;0.001</td>
<td>10.4</td>
<td>0.108</td>
<td>44.0</td>
</tr>
<tr>
<td>Composite</td>
<td></td>
<td>8</td>
<td>3</td>
<td>-0.10</td>
<td>-0.52</td>
<td>0.33</td>
<td>0.655</td>
<td>51.39</td>
<td>&lt;0.001</td>
<td>89.5</td>
</tr>
<tr>
<td>Subsequent Thinking Time</td>
<td></td>
<td>3</td>
<td>4</td>
<td>0.47</td>
<td>0.31</td>
<td>0.64</td>
<td>&lt;0.001</td>
<td>2.1</td>
<td>0.560</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>6</td>
<td>0.39</td>
<td>0.20</td>
<td>0.57</td>
<td>&lt;0.001</td>
<td>6.1</td>
<td>0.299</td>
<td>28.1</td>
</tr>
<tr>
<td>Composite</td>
<td></td>
<td>8</td>
<td>3</td>
<td>0.50</td>
<td>0.32</td>
<td>0.68</td>
<td>&lt;0.001</td>
<td>12.17</td>
<td>0.095</td>
<td>42.5</td>
</tr>
<tr>
<td>Number of perfect solutions</td>
<td></td>
<td>3</td>
<td>3</td>
<td>-0.58</td>
<td>-0.75</td>
<td>-0.40</td>
<td>&lt;0.001</td>
<td>0.2</td>
<td>0.892</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
<td>-1.61</td>
<td>-3.14</td>
<td>-0.08</td>
<td>0.039</td>
<td>38.3</td>
<td>&lt;0.001</td>
<td>98.7</td>
</tr>
<tr>
<td>Composite</td>
<td></td>
<td>8</td>
<td>3</td>
<td>-0.66</td>
<td>-0.85</td>
<td>-0.46</td>
<td>&lt;0.001</td>
<td>13.60</td>
<td>0.059</td>
<td>48.5</td>
</tr>
</tbody>
</table>

Note: SMD denotes the standardised mean difference between groups, Q is Cochrane’s Q and p(Q) its p-value.