Self-confidence and metacognitive processes

Sabina Kleitman*, Lazar Stankov ¹

School of Psychology, The University of Sydney, Sydney, NSW, 2006, Australia

Received 22 May 2006; received in revised form 15 January 2007; accepted 3 March 2007

Abstract

This paper examines the nature of the Self-confidence factor. In particular, we study the relationship between this factor and cognitive, metacognitive, and personality measures. Participants (N=296) were administered a battery of seven cognitive tests that assess three constructs: accuracy, speed, and confidence. Participants were also given the Metacognitive Awareness Inventory (MAI, Schraw, G., and Dennison, R.S. (1994). Assessing metacognitive awareness. Contemporary Educational Psychology, 19, 460–475.), a personality measure of the Big Five factors and our own Memory and Reasoning Competence Inventory (MARCI). Results indicate the presence of separate Self-confidence and Metacognitive processes factors, and a moderate correlation (.41) between them. The Self-confidence factor taps not only processes linked to performance on items that have correct answers, but also sureness level in beliefs about events that may never occur. A hierarchical multiple regression showed that the Self-confidence factor was predicted by accuracy of performance, Metacognitive Awareness Questionnaire, and beliefs of competence in reasoning ability.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Confidence ratings; Metacognition; Self-confidence

1. Introduction

This paper examines the relationship between Self-confidence measured during performance on typical cognitive tests and several conceptually related constructs. The aim is to further our understanding of Self-confidence and establish its status within the taxonomy of cognitive/metacognitive processes.

Four lines of evidence will be considered. First, we shall argue that Self-confidence is a broad psychological trait that cuts across diverse cognitive domains. Second, we shall examine the breadth of the Self-confidence construct and, in particular, whether it can be extended to beliefs about the veracity of predictions of events that may or may not happen in the future. Third, we shall look at the relationship between the Self-confidence factor and measures of broad Self-concepts captured by our own Memory and Reasoning Competence Inventory (MARCI) and the Metacognitive

* Corresponding author.

E-mail address: sabinak@psych.usyd.edu.au (S. Kleitman).

¹ The study reported in this manuscript was conducted at the University of Sydney and supported by the Departmental Research Grant. The data can be obtained by writing to the first author. Lazar Stankov is now working at the Education Testing Service, Princeton, New Jersey.
Awareness Inventory (MAI, Schraw & Dennison, 1994). These measures are expected to show meaningful correlations with Self-confidence, demonstrating that Self-confidence is related to, but cannot be reduced to, these two meta-cognitive domains. Fourth, we shall examine the relationship between Self-confidence, personality dimensions, and measures of Speed of test-taking.

2. The Self-confidence factor

Our procedure for assessing the Self-confidence construct is integrated within the typical test-taking activity. Immediately after responding to an item in a test, participants are asked to give a rating (expressed in terms of percentages) indicating how confident they are that the chosen answer is correct. Accordingly, these ratings directly follow the cognitive act of providing an answer, rather than relying on the participant’s own perceptions of habitual ways of acting. Confidence ratings for all attempted test items are averaged to give an overall confidence score (see Harvey, 1997; Keren, 1991; Stankov, 1999).

There is a substantial amount of data showing individual differences in confidence ratings. The available evidence indicates that confidence ratings are reliable (Stankov, 1998, 1999, 2000). Studies that employ typical cognitive batteries assessing broad factors of fluid (Gf) and crystallized (Gc) intelligence indicate that correlations between accuracy and confidence scores from the same test tend to average between .40 and .60 (see Stankov, 1999, for a review). Thus, they have a meaningful moderate correlation with cognitive abilities — i.e., higher accuracy is linked to a higher confidence level (see Harvey, 1997; Keren, 1991; Stankov, 1999). Nevertheless, numerous studies also show that correlations between confidence ratings from different cognitive tests are high enough to define a broader and more general Self-confidence factor that is separate (yet positively related) to the factors that underlie intelligence (Klayman, Soll, Gonzalez-Vallejo, & Barlas, 1999; Kleitman & Stankov, 2001a; Nietfeld & Schraw, 2002; Pallier et al., 2002; Schraw, Dunkle, Bendixen, & Roedel, 1995; Schraw & Nietfeld, 1998; Stankov, 1998, 1999, 2000; Stankov & Crawford, 1996, 1997). That is, people who are more confident on one cognitive task tend to be more confident across other tasks.

At present, there is considerable information about the relationship between Self-confidence and cognitive abilities (Stankov, 1998; 1999; Stankov & Crawford, 1996, 1997; see also Stanovich, 1999). Something is also known about its relationship to personality (Pallier et al., 2002). Self-confidence is sometimes treated as a personality trait, either on its own or as an underlying facet of broader traits (see Blais, Thompson, & Baranski, 2005, for a review). Some consistent small correlations have been found with the Openness factor from the Big Five model (Pallier et al., 2002). However, it is unclear at present whether this relationship with Openness is mediated by cognitive ability since both Self-confidence and Openness correlate with cognitive ability. Moreover, Blais et al. (2005) have demonstrated that a broad range of cognitive styles, including Need for Cognition and the Desire for Structure, had no effect on confidence.

Importantly, the evidence for the meaningful relationship between Self-confidence and some conceptually related constructs is scarce. In theory, confidence judgments reflect an important aspect of metacognitive processes (see Stankov, 1999, for a review). However, there is a shortage of empirical evidence to justify this claim. Thus, the place of the Self-confidence factor within cognitive/metacognitive taxonomy remains unclear.

2.1. Self-confidence and metacognition

Metacognition refers to one’s awareness of one’s own cognitive processes. It is an assessment of one’s own ability, knowledge, and understanding of task-relevant factors. Metacognition is an essential aspect of information processing, with wide implications for both educational and industrial/organizational settings. Research into metacognitive processes appears to be critical in identifying factors that facilitate intelligent behavior, yet move beyond the (arguably) restricted scope of traditional notions of intelligence (see Bowman, Markham, & Roberts, 2001).

Most theories of metacognition (Nelson & Narens, 1994; Schraw & Dennison, 1994; Schraw & Moshman, 1995) distinguish between knowledge of cognition (i.e., knowledge about one’s own cognitive processes or capabilities, as well as knowledge about how, when, and why to use strategies and allocate cognitive resources), and regulation of cognition (i.e., the control aspect of learning). Three processes of metacognitive regulation are typically posited (Schraw & Moshman, 1995; Schraw et al., 1995): (1) Planning, which refers to the selection of appropriate strategies and allocation of cognitive resources before the task; (2) Monitoring, which refers to the awareness of
understanding and performance during the task; and (3) Evaluation, which refers to the appraisal of performance after task completion.

Confidence scores are said to reflect self-monitoring processes of metacognition (see Keren, 1991; Stankov, 1999, for reviews). However, it is yet to be demonstrated that confidence ratings have meaningful overlap with other conceptually related metacognitive constructs.

3. The relationship between Self-confidence and hitherto unexplored constructs

3.1. Breadth of Self-confidence: answers that cannot be verified

In this paper we extend the measurement of Self-confidence beyond the typical test-taking situation and ask participants to state their level of confidence in the likelihood of events that may or may not take place. In real life, people are often asked to state their opinions about the likelihood of some personal or world event that may or may not happen. These opinions do not have a correct answer at the time of testing and, indeed, correct answers may never become available. Nevertheless, people often indicate how sure they feel about their views. There are some activities – e.g., weather and economic forecasting – that have a similar feel at the time of prediction, but they generate objective outcomes sooner rather than later, or at least eventually. Early work on decision making examined this type of forecasting in relation to the overconfidence bias (see Plous, 1993, for review of this literature). The interest in this topic was fueled, in part, by the assumption that knowledge about the discrepancy between the prediction and the event may play a significant role in the development of forecasting expertise.

To our knowledge, data on the confidence which people have in their judgments about non-verifiable events have never been reported. This data, however, can be useful for understanding the broadness of the self-confidence factor. For this purpose we designed 23 opinion statements. In each statement, participants had to estimate the likelihood (from 0% to 100%) of occurrence of a particular event. The events ranged from those pertinent to Australia (e.g., becoming a republic), to the World in general (e.g., virtual reality becoming the main entertainment in the future), and to themselves (e.g., succeeding in a chosen university course). After stating their beliefs, participants were asked to indicate how sure they were of their opinions on a scale from 1 (Not sure at all) to 5 (Very sure). The mean of these judgments indexed the level of assurance that participants held in their opinions which, for brevity, we shall refer to as the “Sureness” score. Here, however, the question of interest is the relationship between Sureness and typical Self-confidence scores. Are they tapping the same process?

3.2. Metacognitive scales

Questionnaire measures of metacognition exist in the literature. In this study we employ two measures — a well known Metacognitive Awareness Inventory (MAI developed by Schraw & Dennison, 1994) and our own instrument the Memory and Reasoning Competence Inventory (MARCI, Kleitman & Stankov, 2001b). By establishing that scores on questionnaire measures of metacognition show significant correlations with Self-confidence, we can validate the claim that this latter factor indeed taps aspects of metacognitive processes.

3.2.1. The Metacognitive Awareness Inventory (MAI)

The MAI (MAI, Schraw & Dennison, 1994) is a 52-item questionnaire specifically developed to assess people’s self-understanding or awareness of their metacognitive processes. It reflects peoples’ awareness of themselves in a non-specific learning context and it assesses awareness of general learning capabilities and strategies. It measures two major components of metacognition: (1) knowledge about one’s own Cognition; and (2) regulation of one’s own Cognition (Schraw & Dennison, 1994, see also Schraw, 1998). Knowledge about Cognition refers to what people know about themselves as learners, self-understanding of their strong and weak points, strategies, and the conditions under which strategies are most effective (assessed using such items as: “I understand my intellectual strengths and weaknesses”; “I have a specific purpose for each strategy I use”; “I am aware of what strategies I use when I study”). Regulation of Cognition reflects peoples’ perceptions about the ways they plan and implement strategies, monitor and correct

---

2 The opinions were similar to the Probability and Certainty Test (see Brim, 1955). The scale was originally designed to assess direction and the strength of attitudes. For the purposes of this paper, no use of the probabilities assigned to potential events was made.
comprehension errors, and evaluate their learning (e.g., “I ask myself questions about the material before I begin”; “I am a good judge of how well I understand something”; “I know how well I did once I finish a test”). In previous studies, these two domains of metacognition were strongly interrelated, indicating that knowledge and regulation work together to assist in self-regulation (Schraw, 1998; Schraw & Dennison, 1994). Furthermore, the MAI had significant, albeit low, correlations with confidence ratings on a reading comprehension test (Schraw & Dennison, 1994).

3.2.2. Memory and Reasoning Competence Inventory (MARCI)

We designed this scale to assess a particular aspect of the Knowledge about Cognition component of metacognition: knowledge about one’s own memory and reasoning. Moshman (1994) proposed a similar construct labeled as constructive metareasoning. He describes it as “a type of metareasoning that involves the operation of cognition on one’s own reasoning” (p. 141), a sort of “reasoning about reasoning” (p. 141) and, accordingly, reasoning about memory.

The development of MARCI was grounded in the construct of self-concept — a generic term that refers to subjective perceptions of one’s own relative strengths and weaknesses in relation to some general or specific activities. Stankov and Crawford (1997) argued that some well-established areas of academic self-concept might be related to Self-confidence. They suggested that scales for the assessment of Mathematics and English language facets of academic self-concept, such as those used in the Self-Description Questionnaire-II (SDQ, Marsh, 1990), would be correlated with performance on tests of fluid and crystallized intelligence. However, Stankov and Crawford (1997) found only limited support for the hypothesized relationship. Thus, English self-concept had a low (.22) correlation with confidence ratings on a Vocabulary test and Mathematics self-concept had a similar low correlation with confidence ratings based on the Raven’s Progressive Matrices (RPM) Test. Nietfeld and Schraw (2002) reported stronger correlations between a Mathematics self-efficacy questionnaire and confidence ratings assigned to RPM and a Probability Test (correlations ranged from $r = .31$ to $r = .44$, $p < .05$). However, these studies did not control for a possible mediation effect of accuracy of performance on these correlations.

Furthermore, a closer examination of the empirical evidence for the constructs of Self-confidence and Self-concept suggests a major distinction between these in terms of breadth. Self-concepts tend to be domain specific (i.e., limited to a particular domain such as English, Maths, and Physics). It is therefore possible that participants viewed their performance on cognitive tasks such as RPM or Vocabulary tests as being unrelated to their performance in Maths and English. In fact, Ackerman, Beier and Bowen (2002) report that individuals have both generally accurate and differentiated views of their relative standing on abilities and knowledge when these are specifically related to a particular domain. This could explain low correlations between confidence ratings on RPM and measures of academic self-concepts in the Stankov and Crawford (1997) study.

We propose that self-concept measures focusing on activities relevant to cognitive test-taking behaviors should be related to confidence scores. Following on the work of Allwood and Montgomery (1987), the most important strategies employed in typical tests of intelligence are recognition and inference (i.e., memory recall and reasoning). By assessing beliefs in the competency of one’s own memory and reasoning abilities, we can infer people’s knowledge about their cognition. At the same time, employing cognitive measures that assess relevant cognitive abilities should aid to examine veracity of these beliefs. We therefore assumed that self-beliefs in the competency of one’s own memory and reasoning abilities should correlate positively with confidence judgments and performance on tests of intelligence.

To investigate this proposition, MARCI was designed according to the model of self-concept outlined by Marsh and colleagues (see Marsh, Byrne, & Shavelson, 1992). Self-concepts are formed through experience with, and interpretations of, one’s environment. The MARCI incorporates the Internal/External (I/E) Frame of Reference Model proposed by Marsh et al. (1992). The I/E model suggests that English (verbal) and Mathematics self-concepts are distinct because they are formed in relation to both external and internal comparisons. According to the external comparisons principle, the development of Self-concept is influenced by the process of social comparison. A person compares his/her ability in math and reading with the perceived ability of other students in these areas. According to the internal comparisons principle, a person also compares self-perceived ability in maths with his/her self-perceived ability in English. The items of the Memory and Reasoning Competence Inventory (MARCI) reinforced the External (“I can remember more material than the average person”) and Internal (“Compared to my other cognitive abilities [e.g., attention, reasoning], my memory is good”) comparisons (see Appendix A for the items of the MARCI scale).

We expect that beliefs about memory and reasoning competence will be associated with self-confidence on cognitive tests requiring memory and reasoning processes (i.e., crystallized and fluid intelligence markers). We also predict that there will be a meaningful relationship between MARCI facets and the accuracy of performance on these
cognitive tests. However, we expect that the relationship between MARCI and self-confidence will extend above and beyond that of accuracy. Furthermore, by establishing a meaningful relationship between scores on MARCI and MAI questionnaires, we can validate the claim that both measures tap related aspects of the metacognitive processes.

4. Aims of the current study

This study has two aims. Firstly, we wish to establish the status of Self-confidence within the taxonomy of cognitive/metacognitive processes. To achieve this we shall examine its relationships with questionnaire measures of metacognition, in addition to ability, test-taking speed, and personality traits. Secondly, we aim to study the broadness of the Self-confidence factor by including sureness ratings for the unverifiable (currently or ever) events in addition to confidence ratings given to typical cognitive tests.

5. Method

5.1. Participants

The study is based on 296 First Year Psychology students (85 males). Mean age was 19.03 (SD = 3.15). Participants were tested in groups of approximately 20. All participated in the experiment as part of their course requirements. The overall testing time was about 2 hours depending on the individual rate of responding.

5.2. Test Battery

5.2.1. Cognitive Ability Measures

Our battery of cognitive tests was consistent with the fluid and crystallized (Gf/Gc) theory of intelligence (see Horn & Noll, 1994). Tests 1, 3, 6 and 7 are markers of Gf, while tests 4 and 5 are markers of Gc. The nature of test 2 (the Verbal Reasoning Test) suggests that, similarly to the Esoteric Analogies Test, this test may be a factorially complex measure, loading on both Gf and Gc factors.

Tests 2–5 contained confidence judgments. Tests 2–4 were computerized and included measures of reaction time (RT) for each item. Cognitive tests 1, 6, 7 and all questionnaires (measures 8–10 below) were given in a pencil-and-paper format.

1. Quantitative Switching Task (QST) (attention switching measure). Participants were given sixty number strings. For odd items they had to search the string and report the largest even digit. For even items they had to report the smallest odd digit. The test had a 2-minute time limit. For example, item 1: 3 5 6 4 5 2 3 7 (answer = 6); Item 2: 8 7 9 6 5 8 2 2 (answer = 5).

2. Verbal Reasoning Task (VRT) (Kleitman, 2003). The test consists of 30 items. The items were designed specifically to elicit the response selection strategies described by Allwood and Montgomery (1987). To elicit the “immediate recognition” (memory recall) strategy, General Knowledge questions were included. These items covered content areas such as geography, history and lexical knowledge. For example, “Which of the following words means the same thing as ‘facilitate’? (1) Strip; (2) Turn; (3) Help; (4) Bewilder; (5) Intend.” (Answer = 3). Questions were also designed to elicit the “guessing” response strategy. These items were taken from a number of Encyclopedias. The topics covered entailed a very specialized knowledge of biology, statistics, history, and psychology. For example, “What do “moray eels”, “aulopus purpurissatus” and “alabes dorsalis” have in

---

3 This overlap will also reflect veracity of the MARCI scale as it will validate the participants’ claims of competency of their reasoning and memory abilities against their actual performance on the cognitive tests.

4 The original sample included 311 students. Fifteen participants were excluded on the basis of their poor language skills assessed from their biographical information (i.e., first language, years in Australia [if not born in Australia], and frequency of English dictionary usage) and a general impression.

5 The Horn-Cattell theory is a hierarchical model that defines intelligence in terms of independent broad abilities (Carroll, 1993). According to the model, fluid intelligence (Gf) reflects basic abilities in reasoning, while crystallized intelligence (Gc) reflects the effects of acculturation. The model regards Gf and Gc as second-order factors (Horn & Noll, 1994).
common? (1) They all belong to the osteichthyes class; (2) they all come from the family clupeidae; (3) they all come from the family gobiesocidae; (4) they all belong to the asteroidea class; (5) they all come from the ascididae family.” (Answer = 1). The third category of questions was constructed to elicit “inference” (reasoning) strategy. The subject matter of the questions did not contain any information or knowledge content that would be easily accessible from memory. However, hints were built within both the question and the alternatives provided so that participants could reason in order to find a correct response to the question. For example, “A term ‘dyschronaxis’ is used to denote: (1) a superior ability to judge tactile stimulus; (2) an impaired ability to maintain a course of action and/or line of thought; (3) an enhanced sense of smell; (4) an impaired ability to judge what time it is; (5) an impaired ability to walk”. Even though the word “dyschronaxis” would be unfamiliar to the vast majority of people, it is possible to infer the correct meaning of this word by employing “lexical” reasoning – to partition the word into two parts – “dys” – meaning impairment and “chronaxis” – meaning time. Hence, the correct answer is (4).

3. **Nonsense syllogisms.** This measure was taken from the Kit of Reference Tests for Cognitive Factors (French, Ekstrom, & Price, 1963). The test consists of sixteen two-choice items consisting of a syllogism with nonsense content where the participant must decide whether reasoning is valid or not valid. For example, “All trees are fish. All fish are horses. Therefore all trees are horses.” (Answer = valid reasoning).

4. **Esoteric Analogies Test** (Stankov, 1997). The test contains 20 items of the following type: CHICK is to HEN as CALF is to: BULL, COW, COAT, ELEPHANT (answer = COW).

5. The 10-item General Knowledge (GK). This test was taken from the “Gf/Gc Quickie Test Battery” (Stankov, 1997). For example, “Leucocytes are: (1) Small bones in our hands; (2) Blood cells; (3) Small hair cells in our ears; (4) A form of bacteria; (5) Male hormones.” (Answer = 2).

6. **Probabilistic Reasoning Test.** This test was adapted from a High School Mathematics textbook. It included 14 standard probability reasoning tasks, measuring participants’ ability to estimate sample space, and a range of probabilities within this space (Marlin & Nilsson, 1999). For example, “In two child families, what is the probability of there being 2 sons? 2/4; 1/4; 5/6; 3/5; 2/5” (answer = 1/4).

7. **Conditional Reasoning Test.** The test consists of 12 conditional reasoning items (i.e., If p then q). The typical aim of conditional reasoning tasks is to see whether people can make inferences associated with four conditional rules-Modus Ponens (MP), Denial of Antecedent (DA), Affirmation of the Consequent (AC), and Modus Tollens (MT). For an extensive review of conditional reasoning tasks see Manktelow (1999). Participants were asked to evaluate quality (truth or falseness) of the items. The questions’ context was trivial: “If the class is on, there is no noise. There is noise, thus, there is no class” (answer = true, i.e., MT rule).

5.2.2. Other measures

8. “Sureness”. Participants were given 23 opinion statements in which they had to estimate the likelihood (from 0% to 100%) of the occurrence for a variety of events. Participants were also asked to indicate how sure they were about their opinions on a scale from 1 (Not sure at all) to 5 (Very sure). For example, “The probability that a cure for AIDs will be eventually found is about _ in 100.” “How sure you are of your opinion?”

9. **Metacognitive Awareness Inventory (MAI, Schraw & Dennison, 1994).** This 52-item questionnaire assesses two aspects of metacognition: Knowledge about Cognition and Regulation of Cognition. For example, “I ask myself periodically if I’m meeting my goals.” The participants had to evaluate the extent to which each statement described themselves using a 6-point Likert scale ranging from Never to Always.

10. **Memory and Reasoning Competence Inventory (MARCI).** The test consists of 16 items, 8 items for each component. The respondents had to evaluate the extent to which each statement described themselves using a 6-point Likert scale ranging from False to True. Reasoning and Memory items were intermixed.

5.2.3. Personality measure

11. **Openness to Experience (Openness), Conscientiousness, Extraversion, Agreeableness and Neuroticism.** Each scale contains 10 items from the relevant sub-scales of the Christal’s Trait Self-Description Inventory (1993).
6. Procedure

Participants were asked to provide biographical details (e.g., gender, age, first language, years in Australia [if overseas-born]). A timed test—Quantitative Switching—was administered first. To avoid the possibility that performance on cognitive test can influence peoples’ perception of competence of their reasoning and memory abilities, the MARCI scale was administered prior to the administration of the cognitive tests. The remainder of the tests were intermixed and participants instructed to do them in any order they wished.

7. Results

7.1. Descriptive statistics and reliabilities

The descriptive statistics and reliability estimates presented in Table 1 are consistent with previous findings with similar cognitive tests (e.g., Kleitman & Stankov, 2001a,b; Pallier et al., 2002; Stankov, 2000). The reliabilities (Cronbach’s alphas) for the majority of accuracy measures used in this study are somewhat low, yet within acceptable levels for research purposes (Anastasi & Urbina, 1997). In accordance with the previous results, the reliabilities of confidence judgments are consistently high (mean of all alpha coefficients = .89), and they are higher than reliabilities of both accuracy and reaction time measures (means of all alpha coefficients are .57 and .84 respectively; see Kleitman

| Table 1 |
|---|---|---|
| Descriptive statistics and reliability coefficients (Cronbach’s alpha) for all variables |
| Mean | SD | Alpha |
|**Accuracy** | | | |
| Quantitative Switching | 33.43 | 12.87 | N/A |
| Verbal Reasoning | 46.08 | 9.96 | .59 |
| Nonsense Syllogisms | 58.28 | 15.43 | .53 |
| Esoteric Analogies | 63.53 | 16.01 | .66 |
| General Knowledge | 59.00 | 15.68 | .51 |
| Probability Reasoning | 83.99 | 14.04 | .67 |
| Conditional Reasoning | 63.15 | 16.93 | .47 |
|**Confidence** | | | |
| Verbal Reasoning | 42.16 | 11.02 | .91 |
| Nonsense Syllogisms | 75.88 | 11.96 | .95 |
| Esoteric Analogies | 64.89 | 12.04 | .87 |
| General Knowledge | 60.18 | 15.48 | .82 |
| “Sureness” | 3.52 | 0.52 | .89 |
|**Reaction/overall time** | | | |
| Verbal Reasoning | 45.87 | 1.38 | N/A |
| Nonsense Syllogisms | 14.95 | 50.65 | .82 |
| Esoteric Analogies | 11.18 | 30.61 | .86 |
|**Metacognitive measures** | | | |
| MAI | 3.98 | 0.51 | .93 |
| MARCI: memory inventory | 3.87 | 0.97 | .88 |
| MARCI: reasoning inventory | 4.16 | 0.83 | .88 |
|**Personality measures** | | | |
| Openness | 41.80 | 8.70 | .82 |
| Conscientiousness | 45.52 | 8.88 | .87 |
| Extraversion | 47.36 | 8.07 | .84 |
| Agreeableness | 54.75 | 6.93 | .85 |
| Neuroticism | 38.73 | 8.36 | .86 |

Note: Reliability coefficients for the three scores of the Verbal Reasoning Test are based on 16 non-guessing items. The reliability coefficient for the Quantitative Switching task cannot be computed, as it is a speeded test. MAI = Metacognitive Awareness Inventory, MARCI = Memory and Reasoning Competence Inventory.
The reliability estimates for the personality measures are consistent with previous results for this scale and reasonable (ranging from .82 to .87).

The MARCI has high internal consistency (alphas = .88 for both components). This replicates our previous findings (see Kleitman & Stankov, 2001b) and implies that people hold stable beliefs in their memory and reasoning competence. The overall Metacognitive Awareness Inventory (MAI) also had a high reliability estimate of alpha = .93. Internal consistency was also high for the “Sureness” measure (alpha = .89), indicating consistency in the level of ‘sureness’ that people give to their opinions about potential happenings.

7.2. Confirmatory factor analysis: evidence for broad confidence and metacognitive processes factors

To investigate the structure of cognitive and metacognitive measures, a confirmatory factor analysis (CFA) was carried out using the Maximum Likelihood (ML) method from the AMOS program (Arbuckle & Wothke, 1999). Chi-square ($\chi^2$) is one of the most commonly used fit indexes. Small values relative to the degrees of freedom indicate statistically non-significant differences between the actual and the implied matrices, signaling no discrepancy between the hypothesized model and the data. This statistic, however, is sensitive to sample size. Thus, following the current practice we used the root-mean-square error of approximation and its 90% confidence interval (RMSEA) to assess approximate goodness of model fit in the population. Values lower than .05 and a narrower confidence interval suggest a good fit (Hu & Bentler, 1999). We also report the relative likelihood ratio of $\chi^2$ to degrees of freedom ($\chi^2/df$) statistic (values less than 2 are considered to indicate good fit; see Kline, 1998). In addition, we use Goodness of Fit Index (GFI) to reflect a relative amount of covariance accounted by the model, where values .90 and above .95 suggest acceptable and good fits respectively (Hu & Bentler, 1999). Finally, we use the Tucker-Lewis index (TLI), an incremental fit index that has been shown to be relatively independent of sample size (Marsh, Balla, & McDonald, 1988). Values greater than .90 and .95 are considered to reflect acceptable and good fits respectively.

7.2.1. Factor analysis

The models tested were based on findings from previous studies in our laboratory utilizing similar batteries of tests. Previous research (see Stankov, 1999, 2000) suggests that the present study ought to yield four or five factors: (1) Fluid intelligence; (2) Crystallized intelligence (where both factors 1 and 2 are broad abilities that arise from accuracy scores from the relevant cognitive tests; see Method, Section 5); (3) a separate Speed of Test-taking factor (with loadings from all timed scores from tests 2 – 4); (4) a Self-confidence factor (with loadings from all confidence ratings scores, including Sureness measure); and, perhaps, (5) a Metacognitive Processes factor with loadings from MAI and MARCI inventories. The existence of the 5th factor has not been explored before, but we expect that the factor that is defined by these metacognitive measures will emerge, supporting the convergent validity of these measures. Thus, MAI and MARCI inventories should either define the same factor as the confidence judgments (i.e., a very broad Metacognitive factor; the Four-factor Model), or a separate factor, which has a meaningful overlap with the Self-confidence factor, yet separates from it (the Five-factor model).

The Four-factor model was fitted first, with all confidence scores and the Metacognitive Awareness, Memory and Reasoning Inventories defining the forth-very broad Metacognitive Processes-factor. This model has a relatively poor fit: $\chi^2_{122} = 300.87$, $\chi^2/df = 2.5$, the Root Mean Square Error of Approximation (RMSEA) = .07, the Tucker-Lewis index (TLI) = .83, the Goodness-of-Fit Index is (GFI) = .89, and we do not interpret it here.

The Five-factor model was fitted next. In this model, the very broad Metacognitive Processes factor was split into two factors: all confidence scores defining the forth-Self-confidence-factor, and the Metacognitive Awareness, Memory and Reasoning Inventories defining the fifth-Metacognitive Processes-factor.

After several minor modifications (e.g., setting covariances already close to zero to zero and allowing for a few measures to load on more than one factor), a five-factor model produced the most acceptable measures of fit: $\chi^2_{121} = 206.97$; $\chi^2/df = 1.71$, the RMSEA = .05 (its 90% Confidence interval was .04; .06), TLI = .92 and GFI = .93. All the abovementioned indices are within levels that indicate a reasonably good model fit (see Hu & Bentler, 1999). The results of this CFA are presented in Table 2 and interpretation of this model follows.

---

6 Factor analyses did not include personality scales. However, in order to check on the previously reported correlations between Openness and Self-confidence, personality scales are included in the regression analyses reported in the last part of the Results section.
7.2.1.1. Factor 1: Fluid Intelligence (Gf). As hypothesized, this factor is defined by the accuracy scores from Conditional Reasoning, Quantitative Switching, Syllogisms, Esoteric Analogies Probabilistic Reasoning, and Verbal Reasoning tests. This is also a rather broad factor in terms of the cognitive processes captured. However, its Gf nature is rather pronounced as the factor captures such processes as attention switching, inductive, deductive, and numerical reasoning. Notably, the factor also has loadings from the MARCI. The fact that the Reasoning Competency score of the MARCI loads on this factor validates the authenticity of this measure. That is, people’s beliefs about competency of their reasoning abilities correspond closely to their performance on tests that assess different aspects of reasoning. It is worth noting that, somewhat unexpectedly, the “Sureness” measure also has a significant negative loading on this factor. This suggests that high Gf scores imply a lack of readiness to express confidence in predictions of events that may or may not happen, and hence there might be no objective criteria to verify the accuracy of the prediction. Another loading that was not hypothesized a priori was a small, yet significant loading from the Memory facet of the MARCI. This suggests that, similarly to the Reasoning Competency score of the MARCI, peoples’ judgments of competency of their memory share a small, yet meaningful, relationship with the actual performance on different cognitive tests.

7.2.1.2. Factor 2: Crystallized Intelligence/Verbal Reasoning (Gc). As expected, salient loadings on this factor are from the accuracy scores of the General Knowledge, Verbal Reasoning, and Esoteric Analogies tests. Two confidence scores (from the General Knowledge and Verbal Reasoning tests) also have significant loadings on this factor. This too is a rather broad factor, with dominant verbal and learned components.
7.2.1.3. Factor 3: Speed. As expected, this factor is defined by the three test-taking speed measures. Esoteric Analogies accuracy also has a small loading on this factor. This is a Speed of Test-taking factor (see Roberts & Stankov, 1999). There is no hypothesized correlation between Speed of Test-taking and Self-confidence7. In the present paper, the presence of the Speed factor serves the purpose of discriminant validation.

7.2.1.4. Factor 4: Self-confidence. As predicted, this factor is exclusively defined by high loadings of the confidence scores. Notably, the ‘Sureness’ in one’s opinion measure also has a significant loading on this factor, supporting the existence of the broad Self-confidence factor.

7.2.1.5. Factor 5: Metacognitive Processes. As postulated by the Five-factor model, this factor has loadings from the both MARCI facets and the Metacognitive Awareness Inventory. Although not postulated a priori a small, yet significant loading of the ‘Sureness’ measure is psychologically meaningful. That is, this factor clearly taps the metacognitive processes.

7.2.2. Correlations between factors
The bottom part of Table 2 displays correlations between factors. Those correlations that were not significant at the .01 level were fixed at zero. Not surprisingly, two ability factors (Gf and Gc) share a positive correlation ($r = .34$). The Self-confidence factor has substantial (and the highest) correlation with the Metacognitive Processes factor ($r = .41$), and notable correlations with the Gc and Gf factors ($r = .20$ and $r = .34$ respectively). In contrast to the Self-confidence factor, the Metacognitive Processes factor does not correlate with the two ability factors. It should be noted, however, that both facets of the MARCI questionnaire and Sureness measure loaded on both Metacognitive Processes and Gf factors. This implies that there is a relationship between these factors. The Speed factor was orthogonal to all but the Metacognitive Processes factor ($r = .30$).

7.3. Relationship between Self-confidence, Cognitive and Personality Traits
Table 3 summarizes the results of a four-block hierarchical multiple regression analysis with Self-confidence as a criterion and Ability composite, Metacognitive, Speed of Test-taking composites, and five personality dimensions as the predictor blocks8. Although much of the outcome can be anticipated from the results of factor analysis, we

Note: MARCI = Memory and Reasoning Competency Inventory.

Table 3
Model summary of the hierarchical multiple regression analysis with confidence scores as a criterion

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Predictors</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Standard beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Accuracy</td>
<td>.396</td>
<td>.157</td>
<td>.157</td>
<td><strong>.396</strong></td>
</tr>
<tr>
<td>2</td>
<td>Accuracy</td>
<td>.483</td>
<td>.234</td>
<td>.077</td>
<td><strong>.353</strong></td>
</tr>
<tr>
<td></td>
<td>Metacognitive awareness inventory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MARCI: memory inventory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MARCI: reasoning inventory</td>
<td></td>
<td></td>
<td></td>
<td>.192**</td>
</tr>
<tr>
<td>3</td>
<td>Model 2 predictors plus Speed of test-taking</td>
<td>.493</td>
<td>.243</td>
<td>.010</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Model 3 predictors plus Openness Extraversion Agreeableness Neuroticism</td>
<td>.516</td>
<td>.267</td>
<td>.023</td>
<td></td>
</tr>
</tbody>
</table>

Note: MARCI = Memory and Reasoning Competency Inventory.

7 Another type of mental speed – speed of providing answers to tasks of trivial difficulty – has been used in psychophysical studies in conjunction with self-confidence. We do not have a measure of this latter type of mental speed in the present study.

8 Composites were estimated as means of the corresponding variables. That is, Self-confidence composite is a sum of four confidence scores and the Sureness measure; the Speed of test-taking is the average of all corresponding speed measures (see Tables 1 and 2). Performance on all cognitive tests was combined into the Ability component.
use regression analyses to clarify the relationship between self-confidence and other constructs, including hitherto unexamined personality measures.

As can be seen from Model 1, cognitive abilities contribute significantly (with positive beta coefficients) to the overall confidence level, explaining 15.7% of its variance ($\beta=.396; p<.01$). Model 2 shows that after accounting for the abilities composite, the Metacognitive Processes block contributes significantly to the overall level of confidence, explaining additional 7.7% of the variance ($p<.01$). To examine these results further, variables that comprised this composite (both MARCI and MAI scores) were incorporated in the analysis separately. In Models 3 and 4 the Speed factor and all five personality measures were included respectively. Neither block contributed significant additional variance to the prediction of the Self-confidence factor ($\Delta R^2$’s are .010 and .023 respectively, $p>.05$). Thus, we will base our interpretations on Model 2 in Table 3. In this analysis, in addition to the Ability composite, MAI and the Reasoning Competency of the MARCI are the only significant predictors, both with positive betas (.201 and .160 respectively, $p<.01$).

8. Summary and discussion

Metacognitive knowledge and skills are essential components of successful learning since they can guide the choice of strategies, and where necessary, provide for their adjustment (Sternberg, 1997). This paper focused on Self-confidence (regulation of cognition, the control aspect of learning) and certain aspects of knowledge of cognition (i.e., knowledge about one’s own cognitive processes or capabilities) components of metacognition.

Previous research on Self-confidence has demonstrated reliable individual differences in confidence ratings on items from cognitive tests. Much of that work focused on the relationship between Self-confidence and cognitive ability. However, research on the relationship between Self-confidence and other theoretically related constructs has been scarce. The current study clarifies both the nature of the Self-confidence construct and its status within the taxonomy of cognitive/metacognitive processes.

Our results replicate previous findings that Self-confidence is a robust factor which is meaningfully related to, yet is independent of, cognitive abilities. The present study also broadens the assessment of Self-confidence to include measures of sureness in the occurrence of currently unverifiable events. Importantly, the broad Self-confidence factor shares a strong and meaningful relationship with the Metacognitive Processes factor. The Metacognitive Processes factor was defined by both facets of the Memory and Reasoning Competence Inventory (MARCI) and Metacognitive Awareness Inventory (MAI), and to a smaller degree the “Sureness” measure. The MARCI was developed to reflect people’s beliefs about competency in their cognitive abilities (memory and reasoning) captured by typical tests of intelligence, the type of knowledge that Moshman (1994) also classified as constructive metareasoning. The Metacognitive Awareness Inventory (MAI) was designed to reflect metacognitive knowledge about one’s own skills and abilities (knowledge about cognition; see Schraw, 1998). The correlation between Self-confidence and Metacognitive Processes factors was the highest in the present study, clearly placing Self-confidence in the metacognitive realm within the cognitive/metacognitive taxonomy. Furthermore, in accordance with expectations, this study found that Speed of test-taking does not correlate with Self-confidence, demonstrating discriminant validity of the latter construct.

The Reasoning competence component of the MARCI, while defining the Metacognitive Processes factor, also had a salient loading on the Gf factor, pointing to the authenticity of peoples’ beliefs about their reasoning abilities. These findings stress the importance of assessing awareness of one’s own cognitive weaknesses and strengths. This information may help a learner to develop an adequate level of confidence in their cognitive performance and assist in effective use of one’s own cognitive abilities and strategies while learning.

The Memory competence component of the MARCI, while also defining the Metacognitive Processes factor, did not correlate strongly with either the accuracy of performance or Self-confidence. There might be two reasons why the Memory score acted differently in comparison to the Reasoning score. First, it might be that people’s assessment of their own memory is not important for their metacognitive regulation. In fact, it has been demonstrated that people’s beliefs in their memory competence rarely converge with their performance on memory tests (see Hertzog & Robinson, 2005, for a review). The second possibility is that people’s assessment of their own memory is important for

9 In Model 4, when the effect of all variables was assessed simultaneously, abilities and the Reasoning Competence score of the MARCI remained significant predictors of the Self-confidence factor. The MAI lost its predictive power, and the Speed of test-taking showed marginal significance.
metacognitive regulation, but this component of the scale is too general and does not address the complexity of memory ability. In other words, Memory Self-concept fails to make a distinction between, say, *long-term* memory as captured by Tertiary Storage and Retrieval (TSR) and *short-term* memories (SAR) (see Carroll, 1993).

It is also worth noting that while ‘Sureness’ ratings have high reliabilities and overlap with other metacognitive measures (MAI and Self-confidence), they have a negative loading on the fluid intelligence factor. Hence, people with higher fluid reasoning abilities are less ‘Sure’ in their opinions of some future events happening when veracity of these statements cannot be established. This finding may have interesting implications for the study of how rational/objective these judgments are. For example, Stanovich (1999) argues that for a thought process to be considered as being rational, it has to share meaningful positive overlap with cognitive abilities. That is, more intelligent people are expected to demonstrate more complex and more ‘rational’ thought processing in relation to problem-solving.

Appendix A

Items of the Memory and Reasoning Competency Inventory

1. My memory is above average. (M)
2. To solve a problem, I rely on reasoning abilities. (R)
3. Compared to other intellectual abilities (i.e., attention, reasoning) my memory is good. (M)
4. I enjoy being involved in activities that require reasoning of some sort. (R)
5. I can remember more material than the average person. (M)
6. I feel confident when solving problems that require reasoning skills. (R)
7. I’m satisfied with my memory. (M)
8. In an exam situation, I get answers right mostly by reasoning. (R)
9. I describe myself as a person with reasoning abilities that are above average. (R)
10. I rely on my memory to get me through exams. (M)
11. I am happy with my reasoning abilities. (R)
12. I enjoy engaging in activities that require me to remember things. (M)
13. I can reason better than the average person. (R)
14. For exam purposes, I memorize material easily. (M)
15. Compared to my other cognitive abilities, my reasoning is sound. (R)
16. I have good memory. (M)

Note: M = Memory Competency items; R = Reasoning Competency items. The respondents used a 6-point Likert scale ranging from False to True.

References


