Personality and Individual Differences 55 (2013) 915–920

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Assessing mental flexibility with a new word recognition test

Cynthia T. Matthew *, Steven E. Stemler

Wesleyan University, 217 High Street, Middletown, CT 06459, USA

ARTICLE INFO

Article history:
Received 2 March 2013
Received in revised form 7 July 2013
Accepted 17 July 2013
Available online 8 August 2013

Keywords:
Mental flexibility
Cognitive ability
Test development
Behavioral flexibility

ABSTRACT

This paper presents the development and preliminary evaluation of a new word recognition test (WRT) designed to measure individual differences in *mental flexibility*, defined as the ability to solve novel problems in unfamiliar settings. Conceptually designed to simulate problem solving in real world performance situations, the test was developed to recruit fluid and reproductive abilities and the interplay between convergent and divergent thinking. It is based on a framework that integrates and extends previous theoretical and methodological approaches to the study of cognitive ability and creative cognition. The WRT was administered with various cognitive ability and criterion measures to an undergraduate student sample ($n = 266$). Results provide preliminary evidence of construct validity. WRT scores correlated as expected with reference measures of cognitive ability, creative performance, and college performance (GPA). Regression analyses showed the WRT explained an additional 4.5% of variance in college performance over and above traditional cognitive ability measures that take up to five times as long to administer. Results suggest further study is warranted given the potential for its contribution to basic research and applied use.

1. Introduction

The capacity to respond effectively to novel problems and unfamiliar settings is essential to successful performance in a broad range of human activities. The challenge from conceptual and psychometric perspectives is how best to define and measure this capacity. Cognitive ability tests that assess flexible thinking are based on various conceptualizations of fluid intelligence and creative thinking and tend to be lengthy with small to moderate predictive power (Lang, Kersting, Hulsheger, & Lang, 2010; Schmidt & Hunter, 1998). Instruments that are better able to predict individual differences in flexible performance may have practical utility and theoretical value. This paper presents the development and preliminary evaluation of the word recognition test (WRT), a newly designed measure of mental flexibility based on a framework that integrates and extends previous theoretical and methodological approaches to the study of cognitive ability and creative cognition.

The ability to deal with novelty and to adapt ones thinking to new cognitive problems without relying extensively on an explicit base of declarative knowledge has been labeled alternatively *fluid intelligence* (Cattell, 1963, 1971), *analytic intelligence* (Carpenter, Just, & Shell, 1990), and *eductive ability* (Raven, 1952, 2009; Spearman, 1923, 1927.) Measures have traditionally relied on various types of pattern recognition tests (series, classification, matrices and conditions) most notably Raven's advanced progressive matrices (RAPM) (Raven, Raven, & Court, 2003). In contrast, the ability to acquire knowledge, conceptualized as *crystallized intelligence* (Cattell, 1963, 1971) and *reproductive ability* (Raven, 2009; Spearman, 1927) have relied on vocabulary tests like the Mill Hill (Raven, Raven, & Court, 1985).

The study of creative cognition, which is closely related to fluid ability, emphasizes the process of forming novel cognitive structures through unusual associations of elements, combining/reorganizing existing elements in knowledge structures, reframing, and perceiving or utilizing visual imagery (Getzels, 1975; Guilford, 1967; Mednick, 1962; Mumford & Gustafson, 2001).

Some specific cognitive abilities associated with creative cognition include *divergent thinking* or the capacity to generate a variety ("flexibility") and number ("fluency") of ideas (Guilford, 1950, 1967), and *convergent thinking* or the evaluative capacity to identify the best combination of ideas or knowledge elements to produce a "best-fit" or "correct" response (Cropley, 2006). Both divergent and convergent processes ultimately involve forming and modifying perceived patterns. Witkin (1975) field dependence–independence model of cognitive style has also been found to be closely related to fluid ability, and, by extension, creative cognition. The theory distinguishes global thinking (field dependence), the tendency to perceive things as they exist as a whole, from analytic thinking (field independence), and the tendency to impose structure on what is perceived. The embedded figures test (EFT: Witkin, 1950; Witkin, Dyk, Faterson, Goodenough, & Karp, 1962), designed to measure cognitive style, is a figural pattern recognition

* Corresponding author. Tel.: +1 860 304 7152; fax: +1 860 685 2761.
E-mail addresses: cmatthew@wesleyan.edu, ctmatthew@creekrowpartners.com (C.T. Matthew), sstemler@wesleyan.edu (S.E. Stemler).

0191-8869/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.paid.2013.07.464

test, methodologically quite similar to the RPM. The perceptual theory of top-down and bottom-up information processing (Myers, 2012) may be a useful overarching framework to describe both cognitive processing concepts. Convergent and field dependent thinking can be considered more broadly top-down approaches, while divergent and field independent thinking, bottom-up.

Drawing from previous research (Matthew, Beckmann, & Sternberg, 2008; Matthew & Stemler, in press), *mental flexibility* is conceptualized here as the capacity to effectively apply acquired knowledge to new problems through a dynamic process of switching back and forth between convergent and divergent thinking. Traditional tests of fluid ability, made up of figural pattern recognition tasks, assess cognitive ability independent of declarative knowledge. Analysis of the RAPM reveals two distinctive processes: (1) figural, perceptual or gestalt (pattern recognition), and (2) analytic or analogical (capacity to induce various rules) (Carpenter et al., 1990; Mackintosh & Bennett, 2005). In order to better measure the capacity to apply knowledge to solve novel problems in unfamiliar real-world settings, a test design that simultaneously recruits fluid and crystallized abilities and the interplay of bottom-up and top-down information processing strategies seems warranted. Such a measure does not appear in extant literature. Toward this end the WRT was developed.

The WRT is similar to figural tests of pattern recognition commonly used to measure fluid ability and cognitive style; however, it relies on acquired knowledge in the verbal domain. Sentences are presented, in which known words of various lengths are rearranged. Identification of the correct words in each sentence requires simultaneously switching one's attention back and forth between letter, word and sentence to generate possible word alternatives (divergent/analytic/bottom-up) and correctly identify words in each sentence (convergent/global/top down). In this way the WRT is expected to engage alternating top-down/bottom-up processes and ultimately capture the dynamic inter-linkage of fluid and reproductive capacity (Spearman, 1923, 1927), thereby providing a more efficient measure of mental flexibility. Word recognition tests have been developed in the past to measure various conceptualized components of verbal intelligence (word comprehension – reproductive) and verbal fluency (word production – fluid) but not their interplay (Beauducel & Kersting, 2002; French, Ekstrom, & Price, 1963; Thurstone, 1938; Thurstone & Thurstone, 1941).

Based on the foregoing theoretical discussion, we expected that WRT should correlate with measures of general cognitive ability but not too highly. WRT should also correlate with measures of field dependence–independence, as the task engages global (word, sentence) and analytic (letter) cognitive processing. Finally we expect the WRT will correlate with and predict creative and academic performance over and above reference measures. Accordingly, we tested the following hypotheses to assess convergent/discriminant validity (H1), predictive validity (H2), and incremental validity (H3) of the WRT:

H1: WRT scores will correlate positively but not too highly with fluid, reproductive/crystallized ability and field dependence–independence test scores.

H2: WRT scores will correlate positively with creative performance and college performance (GPA).

H3: WRT scores will predict college performance (GPA) over and above traditional measures of fluid and crystallized/reproductive ability.

2. Methods and materials

2.1. Participants

Participants were recruited through fliers and e-mail announcements from three universities in Connecticut, USA during the 2007–2008 academic year. They were told that the purpose of the study was to explore 'mental flexibility' and "how we think outside the box" and were offered $30 for their participation.

Data was gathered from 299 undergraduate student volunteers (mean age = 20 years, SD = 2 years). Approximately 74% of participants were female and 25% were male. The vast majority of participants were native English-speaking (94%, $n = 281$); 6% were non-native English-speaking. In terms of ethnic background, 6% ($n = 19$) were African American, 6% ($n = 17$) were Asian American, 4% ($n = 12$) were Hispanic American, 77% ($n = 229$) were European American, and 7% ($n = 21$) reported 'other'. The average number of years of college completed by the participants was 2.4 years (SD = 1.2).

2.2. Procedure

Data collection was part of a larger study aimed at examining pattern recognition as a basic cognitive process that gives rise to mental flexibility across task domains and required developing new assessment instruments (Matthew & Stemler, in press). Participants engaged in a single, 3 h, paper and pencil testing session conducted at their home university. Testing sessions were broken up into three sections with two breaks; two versions were administered to counterbalance the effect of order on test performance. After participants provided informed consent, a battery of tests was administered by experimenters using a standardized script. Upon completion of the testing session, each participant was paid and provided with a debriefing handout.

2.3. Measures

2.3.1. Word recognition test (WRT; Matthew & Stemler, in press)

WRT design was inspired by jumbled word text circulated on the internet in 2003 as follows:

Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe. (*According to a researcher (sic) at Cambridge University, it doesn't matter in what order the letters in a word are, the only important thing is that the first and last letter be at the right place. The rest can be a total mess and you can still read it without problem. This is because the human mind does not read every letter by itself but the word as a whole*).

While the statement has nothing to do with research conducted at Cambridge University, in a review of relevant research on letter ordering and reading Davis (2003) notes elements of truth in this meme: people can recognize words with their letters rearranged, provided the first and last letter remain unchanged, and identify words of different sizes and degrees of rearrangement with variable degrees of difficulty.

The test developed for this preliminary study was aimed at the English-speaking college student population. Sentences developed by Davis were selected and modified based on their relevance to college student participants. New sentences were written by the senior author, reviewed for relevance by her college student assistant, and rearranged according to Davis's findings. A sample sentence of the WRT is shown in Figure 1.

The WRT requires participants to write down as many words as can be deciphered in a timed session. Various sentences were piloted with a small convenience sample of 15 participants to obtain sentences with variable reported difficulty and time required to complete. Four final sentences were selected in which word length ranged from 4 to 12 letters and the sentence length ratio (mixed up words/total words in sentence) ranged from 6/13