Designing a Creativity Assessment Tool for the Twenty-First Century: Preliminary Results and Insights from Developing a Design-Thinking Based Assessment of Creative Capacity

Grace Hawthorne, Manish Saggar, Eve-Marie Quintin, Nick Bott, Eliza Keinitz, Ning Liu, Yin-Hsuan Chien, Daniel Hong, Adam Royalty, and Allan L. Reiss

Abstract In order to assess a person's creative capacity in real-world situations, we propose a novel Design Thinking Creativity Test (DTCT). The DTCT is based on the design thinking principles and can serve as an assessment that reflects problem-solving needs of the twenty-first century. In particular, the DTCT emphasizes assessment of case-based skills to directly measure an individual's application of creativity during an innovation event. In order to create a robust and standardized creativity assessment, we are currently examining (1) the reliability of DTCT in capturing applied creativity in a simulated real-world setting; (2) the validity of DTCT by relating DTCT scores with other standardized assessments of creativity; and (3) the correlations between neuroimaging data and scores achieved in the DTCT task. In this chapter, we provide a thorough background on already existing tools to assess creative capacity and how our approach advances the current state of the art. We also share challenges faced in collecting and analyzing DTCT data, along with proposed solutions. Lastly, we provide our hypotheses and preliminary insights regarding DTCT's ability to capture applied creative capacity.

G. Hawthorne (✉) • A. Royalty
Hasso Plattner Institute of Design (d.school), Building 550, 416 Escondido Mall, Stanford, CA 94305-3086, USA
e-mail: grace@dschool.stanford.edu; aroyalty@stanford.edu

M. Saggar • E.-M. Quintin • N. Bott • E. Keinitz • N. Liu • Y.-H. Chien • D. Hong • A.L. Reiss
Center for Interdisciplinary Brain Sciences Research, Stanford University School of Medicine, 401 Quarry Road, Stanford, CA 94305-5795, USA
e-mail: saggar@stanford.edu; quintin@stanford.edu; nbott@stanford.edu; ekienitz@stanford.edu; ningl@stanford.edu; yhchien1@stanford.edu; weichenh@stanford.edu; reiss@stanford.edu

H. Plattner et al. (eds.), *Design Thinking Research*, Understanding Innovation,
DOI 10.1007/978-3-319-19641-1_8

1 Introduction

Outside of popular, but dated, psychometric instruments like the Torrance Test of Creative Thinking (TTCT) (Torrance 1990) and Alternative Uses Test (Mendelsohn 1976), a widely applicable creativity measurement that takes into account new information and concepts from recent behavioral- and neuroscience research has yet to emerge from recent creativity studies. While the popular TTCT has picture- and verbal-based exercises to assess an individual's possession of mental creativity characteristics and divergent thinking abilities, it does not address the contemporary construct of creativity and innovation that includes associative abilities and an individual's ability to overtly apply their creativity to real world scenarios.

In order to better assess a person's creative capacity in real-world situations, we developed a novel Design Thinking Creativity Test (DTCT) as a next generation creativity assessment that reflects problem-solving needs of the twenty-first century. In particular, the DTCT emphasizes assessment of case-based skills so as to directly measure an individual's application of creativity during an innovation event.

The Hasso Plattner Institute of Design at Stanford University (Stanford d.school) offers classes specifically aimed at enhancing creative capacity through design thinking skill building. As is custom in the academic tradition, instructors must come up with ways to evaluate their student's progression. Often, students are asked to produce deliverables that are then judged by the instructors and classmates. Although this method has academic value, its purely qualitative nature does not allow a formal assessment or measurement on whether the student's creative capacity has been enhanced. In addition, merely possessing creative capacity is differentiated from having the ability to activate and deploy creative capacity. Thus, it is important to examine whether an individual's creative capacity can be objectively measured in real-world settings and enhanced.

Thus, our primary question is: Can an individual's creative capacity be objectively measured in real-world settings and enhanced, and if so, what essential characteristics should be part of the measurement and how does it align with changes in brain functioning?

The development of a new creativity assessment prompted other questions that will be addressed in this chapter. How reliable is a case-base format in capturing applied creativity in a simulated real-world setting? How do results from test subjects' performance on a creativity assessment correlate with functional neuroimaging data both pre- and post-creativity training? How do the results from the DTCT assessment map to the results from the TTCT assessment? Can a case-based creativity assessment measure the application of creativity skills?

The creation of a statistically robust, well-standardized test relies on the acquisition of large amounts of data during large-scale studies in which representative groups of individuals take the test under standardized conditions (Zucker 2003). In this project, we will align the DTCT with creativity training goals from the design

thinking methodology as our initial attempt to establish face validity. By doing so, we hope to draw connections between creativity training, its effect on an individual's creative capacity, and the potential impact of increased capacity from training. By attaining this goal, we (and others) will be able better evaluate the effects of training/teaching methodologies by utilizing a direct, twenty-first century assessment instrument.

We have already laid preliminary groundwork for a next generation creativity assessment through research efforts from a collaboration between Stanford's Hasso Plattner Institute of Design and Stanford School of Medicine's Center for Interdisciplinary Brain Sciences Research (CIBSR), which were focused on: (a) how creativity is manifested in the brain; (b) if training can enhance creative capacity; (c) what are the associated neural correlates of such enhancement; and (d) whether repeated training is required to maintain the enhancement of creative capacity. This was performed through a rigorous longitudinal study that included two 5-week long interventions/trainings sessions and fMRI brain scans, neurocognitive, and behavioral assessments at three time points (Bott et al. 2014; Hawthorne et al. 2013; Kienitz et al. 2014; Saggar et al. 2014).

2 Assessing Creative Capacity

In this section, we briefly provide a brief definition of creativity as well as thorough background on the previous tools used to assess creative capacity in healthy individuals. These tools are broadly divided into following five categories: (a) convergent thinking based tools; (b) divergent thinking based tools; (c) artistic assessments; (d) self-assessments; and (e) improvisation based tools.

2.1 Creativity in the Twenty-First Century Defined

For the purpose of this study, our definition of creativity is guided by the philosophy of the Stanford d.school while insuring translation to the field of cognitive neuroscience. We define creativity as "*a state of being and adaptation of personal skill sets that enables an individual to synthesize novel connections and express meaningful outcomes*" (Hawthorne et al. 2013). This definition captures the intersection of three different axes. To determine how creative a person, deliverable or process is, these components can be rated along three continuums from—(a) existing to new/novel, (b) linear to synthesizing, and (c) no value/meaning to meaningful. We propose a visual illustration of these continuums with three axes (Fig. 1a). A deliverable or process with high novelty, meaning, and synthesis is considered highly creative and so is the person responsible for this deliverable or process. This person, the deliverable, or the process falls within the upper right and back zone of

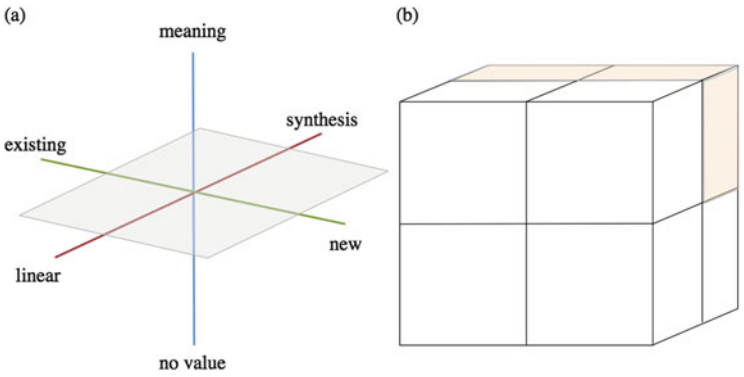

Fig. 1 Visual illustration of our working definition of creativity

the three-dimensional space created by these three axes (the zone in orange in Fig. 1b).

This definition of creativity focuses attention to the person and their skills as opposed to process and outcomes as more traditionally defined. The intention of this focus is to better align the skill of creativity to indications that go beyond the possession of creativity into the ability to exercise/apply it.

2.2 Convergent Thinking Tests

In the extant literature, convergent thinking is defined as the ability to provide the single correct solution to a given problem (Runco 2004). This ability to come up with the right solution, given the constraints of a problem, is thought to require (a) taking a novel approach to the problem; (b) seeing the problem from a different perspective; or (c) making a unique association between the given/associated ideas. The solutions are usually scored in a binary fashion, as either correct or incorrect.

To test the convergent thinking capacity, researchers usually give participants a set of insight (or "aha") problems (Dow and Mayer 2004). The insight problems are different from the broader class of well-structured problems, such that the solution to an insight problem specifically requires a mental set shift or reconceptualization of the problem and given constraints (Goel 2014). Once such mental shift occurs, the problem solver can access the solution with ease. For example, consider the famous insight problem—the triangle (Katona 1940). In this problem, the goal is to arrange six sticks of equal length into four equilateral triangles. Until an individual considers using the third-dimension of space, he/she cannot solve this problem. However, once the shift occurs, and he/she realizes the use of three dimensions, instead of two, the solution presents itself with ease. This phenomenological shift in solving the problem is also referred to as having an "aha" moment.

Another method of assessing convergent thinking capacity is by using the Remote Association Test (RAT; Mednick 1962). In the RAT, participants are presented with three words and are asked to come up with a single word that is associated with all the three given words. For example, when given the three words—"cottage, blue, and mouse", the expected and correct solution is "cheese". As in the case of insight problems, the RAT solutions are also scored in a binary fashion, i.e., as right or wrong.

In summary, convergent thinking tests assess participant's ability to converge to a single correct solution by shifting their mental perspectives. Such thinking, although crucial for creativity, cannot fully capture the creative process.

2.3 Divergent Thinking Tests

Since J. P. Guilford's seminal lecture on motivating research on creativity in 1965, most (if not all) studies have focused on assessing creativity in the form of divergent thinking (Dietrich and Kanso 2010). Divergent thinking is defined as the ability to rapidly generate several ideas/solutions for a given problem. In contrast to convergent thinking, divergent thinking problems are open-ended and do not have one exact solution. Due to the open-endedness, solutions for divergent thinking problems are scored on different putative components of creativity. In the extant literature, some of the widely used components are—(a) originality, i.e., statistical infrequency of the response); (b) fluency, i.e., number of responses; (c) flexibility, i.e., number of unique classes of responses; and (d) elaboration, i.e., the amount of detail in the response.

One of the most widely used tests of divergent thinking is the Torrance Test of Creative Thinking (TTCT; Torrance 1990). The TTCT is available in both verbal and figural versions. The TTCT figural is more widely used as it is less influenced by vocabulary and is appropriate at all levels of age (kindergarten through adult). In TTCT-Figural, participants are given an incomplete figure and are asked to complete the figure in a given amount of time. Participants are also told to complete the figure in such a way that it would portray the most "unusual and unique" story. Trained staff members at the Scholastic Testing Services (STS) later rate the final drawings.

Other divergent thinking tests include Wallach and Kogan's assessments of creative capacity (Wallach and Kogan 1965). In this test participants are asked come up with as many possible items as they can in specific amount of time for a given component. For example, if the given component is "wheels", participants are asked to come up with as many items as they can which contain the component wheels. Like the TTCT, Wallach and Kogan's test is also rated on flexibility, fluency, originality, and elaboration sub-scales.

Other examples of divergent thinking tasks include Alternate Uses Test (AUT; Guilford 1950), where participants are asked to come up with as many unusual or

alternate uses of a given item in fixed amount of time. For example, participants could be asked to rapidly provide alternate uses for a "pen".

In summary, divergent thinking tests are designed to assess a participant's ability to rapidly respond and produce as many unusual or original ideas as possible. It is important to note, however, that the focus in divergent thinking tests is on quantity and not quality/utility of ideas. Thus, divergent thinking is focused on a specific, albeit crucial, idea generation aspect of creativity.

2.4 Artistic Assessments

In addition to testing the creativity process, via convergent and divergent thinking, researchers have also focused on assessing the end product for its creative content. Such assessments are commonly done for artistic products, e.g., paintings, story, poem etc. Two or more field experts typically perform evaluations of such artistic products.

As artistic assessments are limited to professional artists or creative individuals, the Barron-Welsh Art Scale (BWAS) was developed to capture similar assessments in nonprofessional artists (Barron and Welsh 1952). The BWAS is a figure preference test, which includes simple and complex figures that are liked and/or disliked by established professional artists. The participant is shown these figures and is asked to give his/her preference about their liking/dis-liking of the figure. An individual scores higher on BWAS if his/her liking is inline with the artist's preference (Gough et al. 1979).

Although BWAS scores of creativity have been shown to be associated with other indicators of creativity (Gough et al. 1979), there is strong controversy surrounding BWAS' validity and the specific dimension of creativity assessed (Ridley 1979).

2.5 Self-Assessments

In addition to assessing creative process and product, researchers have also investigated different aspects of the creative person him/herself using self-assessments. Such self-assessments measure a participant's creative confidence in generating novel ideas and employing out-of-the-box thinking. Several self-assessment scales were developed over the last 50 years, including Runco's ideation and behavior scale (Runco et al. 2001), the Creative Attitude Survey (Schaefer and Bridges 1970), the Creative Achievement Questionnaire (Carson et al. 2005), and more recently, the Creative Confidence Scale (Royalty et al. 2014).

Self-assessment scales provide a unique indicator for "personal" aspects of creative thinking. However, such scales by themselves are not enough to measure

creative capacity and are generally used in conjunction with other measures of creative thinking.

3 Designing a Design Thinking Based Creativity Assessment

To design the next generation assessment of creativity, we utilized the conceptualization of creativity as prescribed by the Stanford d.school. Further, we narrowed down to specific attributes or characteristics that are critical to creativity and innovation in the twenty-first century. Next, we looked at the landscape of other existing creativity assessments to determine applicability and deficiencies as related to our measurement objective. By determining which cognitive and psychological attributes are measured by existing assessments and whether they addressed our definition of creativity, we saw an unaddressed opportunity to assess creative skill activation versus possession. Existing assessments largely overlook applied or real-world creative acumen that allow participants to demonstrate basic design thinking practices including need finding, creating a point of view, ideating, prototyping, and iterating within constraints or in the face of changing conditions. Thus, an important goal in developing the novel DTCT assessment was to fill these gaps and develop a creativity assessment to measure applied and real-world creative acumen. In this section, we provide details about DTCT's design and its components.

3.1 Case Based Assessment

A case based approach that allows participants to deploy creative skills with intention opens up the potential of assessing creative capacity in a manner that goes beyond the domains of divergent or convergent thinking. By including the creation and building of a real solution from a provided set of cues within a given set of constraints, participants have the opportunity to create a deliverable with high novelty, meaning, and synthesis within an assessment framework. But similar to existing divergent and convergent thinking assessments, in this case based approach, participants also have the opportunity to both conceptualize numerous possible solutions and narrow on one particular outcome of their choosing during the assessment. The case based approach that we have created is a next generation hybrid that builds on pertinent aspects of previous creativity assessments. Our case based approach focuses on assessing applied thinking and adaptability of solutions when faced with new constraints.

Fig. 2 DTCT assessment steps and description

STEP	ACTIVITY
1	Condition evaluation
2	Condition identification
3	Need distillation
4	Point of view ideation
5	Point of view distillation
6	Prototyping
7	Expression/communication with words
8	Expression/communication with pictures
9	Condition change and recast condition
10	Prototyping
11	Point of view distillation
12	Debrief

3.2 DTCT Components

The DTCT was developed around Stanford d.school's design project structure, a fast-paced approach that runs the participant through a full design cycle. Instead of pairing up with another person in a typical design project, participants administered the DTCT run through a course of prompts and activities to create a design solution that is "useful and meaningful" to a hypothetical user.

In the course of 30 min, participants employ characteristics of applied thinking and adaptability for need-finding as well as developing novel solutions. The "case study" assessment begins with a photographic image that consists of people in an environment. Three versions of the DTCT were created for up to three different assessment points; all have the same twelve steps and prompts, and vary in the opening environment and related condition change (Fig. 2).

4 Administration and Data Collection

We prototyped and administered the DTCT as part of another study focused on individual creative capacity building (Hawthorne et al. 2013). Specifically, the DTCT was administered along with several other assessments and MRI scans to 36 participants at three different time points during the 12-week study.

The challenging aspect of administering this assessment is that the activities are open-ended and involve several stages of need finding and brainstorming, which lead to couple steps requiring the development of three dimensional prototypes. To develop 3-D prototypes, all participants were provided with a similar set of limited supplies (e.g., tongue depressors, tin foil, pipe cleaners, post-its, etc.). To capture

and later on rate each prototype, photographs of the prototypes were taken after completion of the respective prototyping steps. In our prototype, each study participant did the DTCT on three different occasions over three measurement periods during a 12 weeks period. Because the structure of the assessments were similar, multiple versions could be administered simultaneously.

A large amount of multi-dimensional data already collected from the aforementioned creative capacity study will allow us to refine and improve the DTCT by investigating associations between different domains of the DTCT and other measures of creativity, cognition, personality and brain function.

4.1 Assessment Timing

We collected data related to creativity, cognition, behavior and brain function before (time 1 or T1) and after (time 2 or T2) a creative capacity building intervention based on the Creative Gym class offered at Stanford's d.school (link: http://dschool.stanford.edu/creative-gym-a-design-thinking-skills-studio/). We included a parallel control group receiving a Chinese language and character drawing learning intervention ("Language Gym") as a non-creativity enhancing control condition. Each group "Crossed-over" to the alternate intervention followed by behavioral assessments and brain scans at time 3 (or T3). Therefore, the DTCT was administered to both groups at three different time points (T1, T2, T3) (Fig. 3).

4.2 Assessment Conditions

The assessment was administered by two facilitators in a group setting in an open room. Each participant was assigned a number and given a version of the DTCT activity booklet that they had not previously received. Each step had an assigned time allotment that was monitored by the facilitators, who were available to answer any questions and capture photographs of the 3-D prototypes built by participants. Photos of the built prototypes had to be accurately labeled during the assessment to ensure the images could be properly matched to a respective participant's assessment booklet.

5 Data Analysis

The data collected during each DTCT assessment include photographs of the prototypes developed, as well as in process thoughts and designs scribbled on the paper booklet. After de-identifying and scanning the booklets and photographs, a web-based assessment tool was generated for each participant. Using this

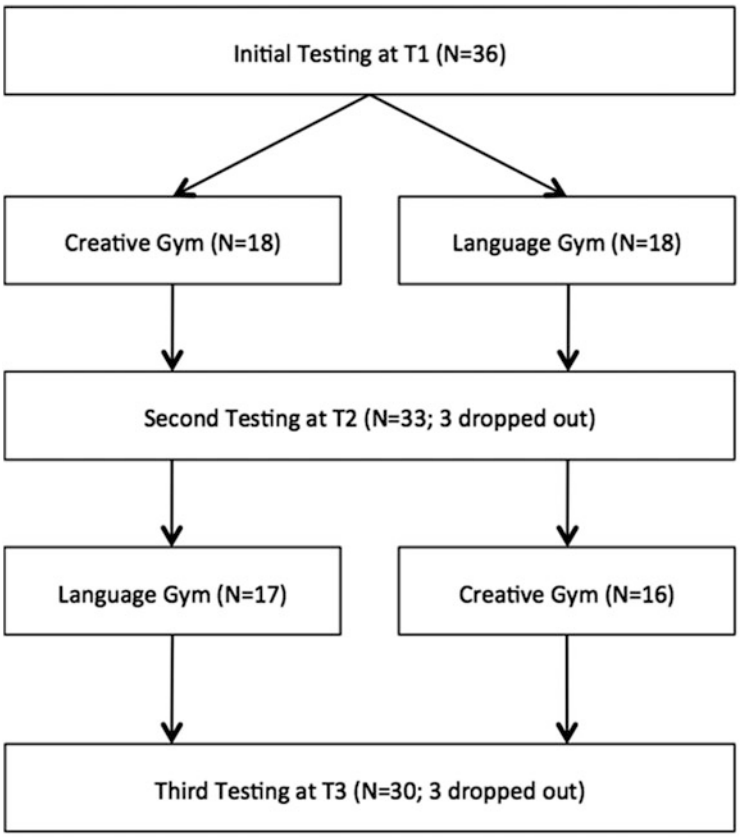

Fig. 3 Visual overview of DTCT administration timing

web-based tool, the raters (at least two) assess and evaluate the work in an efficient, reliable and economic fashion. To validate the novel DTCT assessment, we are currently conducting the following three analyses.

5.1 Convergent Validity

Convergent validity analysis (CVA) will be performed to examine the degree to which the novel DTCT assessment is theoretically related to other standardized and established assessments of creativity (e.g., TTCT). The CVA is usually estimated by correlating scores from two different assessment tools. High CVA scores generally indicate a successful convergence and thereby provide evidence for measuring same/similar underlying concept using the novel and previously established tools. For CVA we plan to use the time 1 data, i.e., data from all 36 participants before intervention.

5.2 Test–Retest Reliability

The reliability of the DTCT over time is an important component of establishing the overall validity of the tool. We will utilize the individual DTCT scores from participants before and after they receive the Language Gym to perform test–retest reliability. We hypothesize that 5 weeks of Language Gym should not significantly affect an individual's creative capacity. Thus, high correlation values between DTCT scores before and after Language Gym are expected.

5.3 Discriminative Validity Analysis

To examine how much value DTCT adds above and beyond the standardized tests of creativity (e.g., TTCT), discriminative validity analysis (DVA) will be performed. In this approach, we hope to show that the DTCT provides additional information about the cognitive skills associated with creativity as well as intervention related outcomes. For DVA, we will use measures of executive functioning, intelligence, personality, as well brain imaging data.

6 Establish a Scoring Guide and Norms

The standardized administration, scoring procedures and norms, and the development and evaluation of the DTCT will begin with already administered tests. Given the objectives and steps included on the test, norm-referenced measures of creative capacity and application will be scored by a host of categories derivative of the TTCT's widely used scoring workbook (Ball and Torrance 1984). The categories will be constructed around key design thinking attributes in the assessment. We will ask the Stanford d.school's most gifted design thinkers to take and score the test to establish the highest possible composite score. We will apply the prototype scoring guide and norm to the already administered tests from the creativity capacity study (108 tests) and to the wider test population that will take the DTCT for this study to refine reliability and validity of the assessment.

7 Looking Ahead

Wider testing of the DTCT prototype will be executed to refine assessment format and improve simplicity and test structure. This will include assessment of the reproducibility of the instrument by administration of alternate forms of the DTCT to subjects within a proscribed period of time. In addition, full analysis of

newly administered DTCT test scores to additional subjects will be performed and correlated with the subject's TTCT scores before, during, and after a time period that will include creativity training.

By creating a psychometrically robust creativity assessment measure as a companion to the TTCT, we will be able to create an assessment tool that can map creativity training to development and real world practice to impact. These findings will help guide instructional content and training exercises for creativity training/intervention in a large number of settings. Increasing creative capacity in individuals goes beyond classroom methodology to applied execution in real world scenarios for impactful change. The DTCT has the potential to become a new industry assessment norm for business professionals, educators and executives across all disciplines and industries as the creativity assessment that matches problem solving needs of the twenty-first century.

7.1 Wider DTCT Administration and Further Refinement

In order to reduce potential confounding covariates from the original prototype and increase statistical robustness, we will administer the DTCT to a larger sample of subjects for this study. We will correlate quality and quantity long-term data from the subjects that participated in the creativity capacity building project, namely information from both neuroimaging and creativity assessments will be analyzed. We will administer the DTCT to two-three additional graduate-level courses at Stanford's d.school to further refine the test language, to create a test administration manual, and to strengthen the establishment of a norm group for reliability and validity. The results from the wider administration of the DTCT will also help us determine the effectiveness of the case-driven hybrid format as real-world scenario.

References

Ball OE, Torrance EP (1984) Torrance tests of creative thinking streamlined scoring workbook, figural A. Scholastic Testing Service, Bensenville, IL

Barron F, Welsh GS (1952) Artistic perception as a possible factor in personality style: its measurement by a figure preference test. J Psychol 33(2):199–203. doi:10.1080/00223980.1952.9712830

Bott N, Quintin E-M, Saggar M, Kienitz E, Royalty A, Hong DW-C et al (2014) Creativity training enhances goal-directed attention and information processing. Think Skills Creat 13:120–128. doi:10.1016/j.tsc.2014.03.005

Carson SH, Peterson JB, Higgins DM (2005) Reliability, validity, and factor structure of the creative achievement questionnaire. Creat Res J 17(1):37–50. doi:10.1207/s15326934crj1701_4

Dietrich A, Kanso R (2010) A review of EEG, ERP, and neuroimaging studies of creativity and insight. Psychol Bull 136(5):822–848. doi:10.1037/a0019749

Dow GT, Mayer RE (2004) Teaching students to solve insight problems: evidence for domain specificity in creativity training. Creat Res J 16(4):389–398. doi:10.1080/10400410409534550

Goel V (2014) Creative brains: designing in the real world. Front Hum Neurosci 8:241. doi:10.3389/fnhum.2014.00241

Gough HG (1979) A creative personality scale for the adjective check list. J Pers Soc Psychol 37:1398–1405

Guilford JP (1950) Creativity. Am Psychol 5(9):444–454

Hawthorne G, Quintin E-M, Saggar M, Bott N, Keinitz E, Liu N (2013) Impact and sustainability of creative capacity building: the cognitive, behavioral, and neural correlates of increasing creative capacity. Springer, Cham, pp 65–77. doi:10.1007/978-3-319-01303-9_5

Katona G (1940) Organizing and memorizing: Studies in the psychology of learning and teaching. Morningside Heights, Columbia University Press, New York

Kienitz E, Quintin E-M, Saggar M, Bott NT, Royalty A, Hong DW-C et al (2014) Targeted intervention to increase creative capacity and performance: a randomized controlled pilot study. Think Skills Creat 13:57–66

Mednick S (1962) The associative basis of the creative process. Psychol Rev 69(3):220. doi:10.1037/h0048850

Mendelsohn GA (1976) Associative and attentional processes in creative performance1. J Pers 44 (2):341–369. doi:10.1111/j.1467-6494.1976.tb00127.x

Ridley DR (1979) Barron–Welsh scores and creativity: a second look. Percept Mot Skills 49(3):756–8

Royalty A, Lindsay NO, Bernard R (2014) Acting with creative confidence: developing a creative agency assessment tool. In: Plattner H, Meinel C, Leifer L (eds) Design thinking research. Springer International Publishing, Heidelberg, pp 79–96

Runco MA (2004) Creativity. Annu Rev Psychol 55:657–687. doi:10.1146/annurev.psych.55.090902.141502

Runco MA, Plucker JA, Lim W (2001) Development and psychometric integrity of a measure of ideational behavior. Creat Res J 13(3–4):393–400. doi:10.1207/S15326934CRJ1334_16

Saggar M, Hawthorne G, Quintin E-M, Kienitz E, Bott NT, Hong D et al (2014) Developing novel methods to assess long-term sustainability of creative capacity building and applied creativity. In: Plattner H, Meinel C, Leifer L (eds) Design thinking research. Springer International Publishing, Cham, pp 29–39, doi:10.1007/978-3-319-06823-7_3

Schaefer CE, Bridges CI (1970) Development of a creativity attitude survey for children. Percept Mot Skills 31(3):861–862. doi:10.2466/pms.1970.31.3.861

Torrance EP (1990) Torrance tests of creative thinking. Figural forms A and B. Scholastic Testing Service, Bensenville, IL

Wallach MA, Kogan N (1965) Modes of thinking in young children. Holt, Rinehart & Winston, New York

Zucker S (2003) Fundamentals of standardized testing. Pearson Inc, New York