A comprehensive method for the measurement of everyday creativity

Guillaume Fürst & François Grin [1]

[1]University of Geneva, Switzerland

Correspondence concerning this article should be addressed to

Guillaume Fürst, Université de Genève

Boulevard du Pont d'Arve, 40, CH-1211 Genève 4.

E-mail : guillaume.fuerst@unige.ch Phone: (+41-22) 379 85 58.

This research received funding from the European Community's Seventh Framework Programme under grant agreement No. 613344 (Project MIME).

Abstract

This paper introduces a multivariate and theoretically driven measurement scheme for everyday creativity. We start by briefly reviewing classical distinctions (e.g., creative person, process and product; potential and manifest creativity) and examining several instruments for the assessment of creativity. We then propose a measurement method that integrates these theoretical elements and psychometric traditions. This method includes several questionnaires (assessing aspects of personality, cognitive styles, creative interests, activities, and achievements) and creativity tasks (divergent thinking, insight, and "real-life" creativity tasks). The scoring method of these tasks rests both on classical indices (e.g., fluency) and on an innovative subjective scoring procedure. This procedure relies on a randomized, double-blind peer evaluation, where some participants rate the creativity of other participants, and allows for unprecedented levels of efficiency and practicality. Overall, results shows that our measurement scheme is psychometrically and theoretically sound. These results are discussed at a theoretical and methodological level, and avenues for future research are explored.

Keywords

Creativity assessment; divergent thinking; subjective scoring; openness; creative potential; creative interests, activities and achievements

A comprehensive method for the measurement of everyday creativity

Creativity is a vast and complex topic; both its conceptualization and measurement raise important questions. In this paper, we focus on the challenges revolving around the measurement of everyday creativity. We start by delineating everyday creativity and reviewing classical distinctions in the field. We then move on to a review of the existing methods available for the measurement of everyday creativity. This review yields an in-depth assessment of these methods, in terms of conceptual focus, psychometrics properties, and practical advantages and drawbacks. We end the theoretical part of the paper by proposing a comprehensive and realistic measurement scheme. The empirical part provides evidence in favor of this method.

Circumscribing everyday creativity

A first important conceptual question concerns the levels of creativity. Historically, creativity researchers have typically distinguished between *small-c* or everyday creativity (creative hobbies, problem-solving in leisure or work activities) and *Big-C* or eminent creativity (high-level creativity with major impact on culture). Beghetto and Kaufman (2007; Kaufman & Beghetto, 2009) have proposed a finer distinction in four levels: *mini-c*, which represents simple combinations of basic pieces of information, involved for instance in learning; *little-c* (or *small-c*), which is equivalent to everyday creativity; *Pro-C*, which refers to progression beyond little-c found in professional-level creative activities (e.g. typically art and science, but in any other domain as well); and *Big-C*, which is equivalent to eminent creativity (national- and international-level creativity). In this paper, we essentially focus on little-c and, to some extent, Pro-c.

Another important distinction is the one between creative potential and creative realization or achievement (e.g., Runco, 2013; Sternberg, Grigorenko, & Singer, 2004). Creative potential usually refers to the presence, in the individual, of various personality traits

(e.g., openness) and aptitudes (e.g., divergent thinking) known to be relevant to creativity. The notion of creative realization or achievement refers to manifest, observable creativity, either in the form of a single product or as the overall creative output of a given person at a given point in time. Understanding how potential and realization are connected is virtually equivalent to understanding how people can move from mini-c to little-c to Pro-C to Big-C (see Kaufman & Beghetto, 2009). Although we will not go into much detail about such transformations in this paper, the instruments we propose encompass both creative potential and creative achievement.

Another essential distinction is the one between domains, for instance between art and science. In principle, such distinctions can be almost infinitely refined (e.g., visual arts, photography, landscape photography, etc.). The question of domain is also connected to another difficult one: is creativity domain-general or domain-specific? Although there is controversy over this issue (e.g., Plucker & Beghetto, 2004), creativity is probably both domain-general and domain-specific, possibly arranged in a hierarchy of micro domains nested in domains grouped within a general thematic area (Kaufman & Baer, 2005). Some traits and aptitudes are arguably general and relevant to any domains (e.g., openness), whereas others are more useful for certain domains only (e.g., extraversion for performance arts). It is also likely that for little-c, domain specificity is rather low, but it is extremely high for Big-C (Kaufman & Beghetto, 2009; Kaufman, Beghetto, Baer, & Ivcevic, 2010). Although this question of domain is peripheral to this paper, we share the view that several domains should be represented, in creativity tasks and/or in creativity questionnaires.

Finally, the distinction between the creative person, process, and product may be the most established one in creativity research (e.g., Rhodes, 1961; Runco, 2004; see also Glăveanu, 2013). The creative *person* refers to all individual variables potentially related to creativity, encompassing many cognitive abilities, personality traits or emotional states. The

creative *process* refers to the dynamic progression of thoughts and actions that a person deploys in order to achieve a creative *product*, that is, the final, observable output[1]. The fundamental person-process-product distinction is of particular interest to this paper, because a method claiming to assess creativity comprehensively should be able to assess, one way or another, several – or better yet, all – of these components. We return to this question when introducing our synthetic approach. Before that, let us review the available instruments allowing the assessment of various facets of everyday creativity.

Measures of everyday creativity

Historically, creativity has been approached through a very large number of methods. It is practical to arrange them in two main categories, namely, questionnaire-based and task-based methods. Each of the methods reviewed below is discussed with reference to the distinctions mentioned above.

Questionnaire-based measures

Questionnaire-based measures of creativity have long been used to assess several facets of creativity, from person to process to product. However, those questionnaires are chiefly used to measure three main sets of variables: (i) personality traits; (ii) thinking style and (iii) creative activities and achievements. Together, these tools can offer a good estimation of individual creativity, whether potential or manifest.

Personality inventories. Several personality factors can be seen as indirect indicators of creativity, allowing for a rough estimate of creative potential. Classically, these are factors from the Big 5 (e.g., openness; McCrae, 1987) or some other general model (e.g., Psychoticism; Eysenck, 1995), as well as more specific traits (e.g., novelty seeking; Schweizer, 2006) or heterogeneous adjective check lists combining a large variety of traits

[1] In addition to these three components, creativity researchers often consider a fourth one, the environment (e.g., various constraints, opportunities, and contextual elements).

(e.g., Gough, 1979). This literature is vast and has a long history, and different lines of work in relatively recent past research allows for a synthesis (e.g., Fürst, Ghisletta, & Lubart, 2016; Fürst & Lubart, 2017). However, unless one has a specific interest in creativity and personality, a measure of openness is probably sufficient to cover this ground. There is indeed overwhelming evidence that openness is *the* personality factor related to creativity (Oleynick et al., 2017). Other personality factors are less consistently associated with creativity; they are relevant only for some specific facets, domains or levels of creativity (e.g., Batey & Furnham, 2006).

Openness/Intellect is a rich and complex personality factor. For this reason, and especially in the context of creativity research where it is so important, two main aspects of this factor can be distinguished: *openness* (encompassing even more specific facets such as interests in aesthetics, fantasy, imagination and reflection) and *intellect* (encompassing specific facets such as interest in truth and ideas, as well as intellectual engagement, quickness, and competence) (DeYoung, Quilty, & Peterson, 2007)[2]. Studies have shown that the intellect aspect is more closely related to intelligence and working memory, while openness is more closely related to schizotypy and apophenia – the tendency to perceive patterns or causal connections where none exists (DeYoung, Grazioplene, & Peterson, 2012; DeYoung, Shamosh, Green, Braver, & Gray, 2009). Furthermore, the Intellect aspect seems more related to science and the Openness aspect closer to the arts (Perrine & Brodersen, 2005). However, the relation between the overall Openness/Intellect factor and creativity clearly appears to transcend all domains (Oleynick et al., 2017).

Thinking styles inventories. Thinking styles (or cognitive styles) are at the crossroads between personality traits and cognitive processes; they can be defined as individual

[2] We use the terminology and writing conventions proposed by DeYoung et al.: *Openness/Intellect* refers to the main factor, while the terms *Openness* and *Intellect* refers to the two aspects of the overall factor.

preferences in thinking, as well as in perceiving and remembering information (Kozhevnikov, 2007). Just as for personality traits, there are numerous thinking styles, and it is beyond the scope of this paper to review them all. The list of thinking styles potentially relevant for creativity is also quite long – see for instance the *innovator* and *adaptor* styles (Kirton, 1976), the *divergent* and *convergent* thinkers (Brophy, 2000), or the various thinking styles proposed by Sternberg (1997).

The Generation/Selection distinction proposed by Fürst and colleagues (Fürst, Ghisletta, & Lubart, 2012; Fürst et al., 2016; Fürst, Ghisletta, & Lubart, 2017) can provide a synthesis. Generation encompasses preferences for the production of many ideas and remote associations, based on fast, automatic, shallow and heuristic processes. Selection encompasses preferences for the careful analysis and evaluation of ideas, based on slow, controlled, deep and systematic processes. According to this model, both Generation and Selection are relevant to creativity. Although most people have a preference for one or the other, both contribute to creativity. Moreover, even if the "content" or the specific ideas subject to Generation and Selection processes obviously vary from domain to domain, the overall Generation and Selection mechanics is hypothesized to be crucially important in any domain.

This approach is not strictly a model of cognitive style *per se*, but it can nonetheless serve as a unifying framework, because Generation is very close to the innovator style and divergent thinker respectively described by Kirton (1976) and Brophy (2000), while Selection is very similar to the adaptor style and convergent thinker. Moreover, Generation and Selection are also closely related to personality factors: Generation is positively correlated to openness, extraversion, non-conformism and impulsivity; Selection is correlated to achievement striving and persistence (Fürst et al., 2016). Hence, to a certain extent, these variables also carry information about personality.

Self-reported creative activities and achievements. Measures based on self-reports of creative activities and achievements are clearly of a different nature than the personality traits and thinking styles described above. While traits and styles represent somewhat indirect indicators of creativity, activities and achievements questionnaires are more direct: they explicitly ask respondents to indicate their actual activities and achievements in various domains. These instruments clearly assess manifest creativity. Some assess many everyday creative activities with relatively basic questions; others consider a more limited number of domains and offer a more in-depth, multi-item measurement.

For instance, The *Biographical Inventory of Creative Behaviors* (BICB; Batey, 2007), consisting of 34 yes/no items, targets very different creative behaviors (compose a poem or a piece of music, start an association, draw a cartoon, produce a theory, design and plant a garden, etc.). The *Creative Behavior Inventory* (CBI; Dollinger, 2003, Hocevar, 1979) follows a similar logic. Another comparable, yet little-known measure is the one used by Verhaeghen, Joorman, & Khan (2005). In terms of content, this questionnaire is similar to those just mentioned, but the answering scheme is more subtle; participants are asked to rate separately the *seriousness* of the practice and the *time spent* on the target activities.

Other questionnaires assess people's beliefs about their level of creativity, again in several domains. Such questionnaires include for instance the *Kaufman Domains of Creativity Scale* (K-DOCS; Kaufman, 2012; McKay, Karwowski, & Kaufman, 2017) or the *Creativity Domain Questionnaire* (CDQ; Kaufman, Cole, & Baer, 2009). One important and interesting feature of these questionnaires is that they propose an explicit distinction between domains. The items regroup in different factors, unlike other measures mentioned until now, which are generally summed up into a single, overall score. One drawback, however, is that these instruments do not directly assess creative activities and achievements.

Finally, a last type of instrument focuses on high-achievement creativity in a selected number of artistic and scientific domains. A well-known instance of such questionnaires is the Creative Achievement Questionnaire (CAQ; Carson, Peterson, & Higgins, 2005) which assesses creativity in ten domains, using specific questions for each domain. These questions have been developed in collaboration with experts in each domain, and so has the relatively sophisticated scoring scheme (the total score for each domain is computed in both an additive and a multiplicative fashion). This questionnaire is better suited for Big-C or Pro-C; in the general population, it yields non-normal, positively skewed and zero-inflated distributions.

In sum, many questionnaires tapping into different aspects of creativity are available. Although these instruments had a rather bad reputation for a long time, recent studies suggest that they actually are quite reliable and valid (Silvia, Wigert, Reiter-Palmon, & Kaufman, 2012). Nevertheless, they can be seen as relatively weak measures of creativity because they are based on self-evaluation only. Measures based on actual creativity tasks can compensate for these limitations. We turn to these now.

Task-based measures

Measures based on lab tasks and direct evaluation of actual creative products can roughly be divided in three general categories: (i) divergent thinking tasks; (ii) artistic and real-life creativity tasks; (iii) insight tasks (Dietrich & Kanso, 2010). As for the questionnaire-based measures, we review these below and put them in relation with the synthesis provided at the beginning of the paper.

Divergent thinking (DT) tasks. DT tasks are rooted in intelligence testing (e.g., Guilford, 1950, 1956) and classically use an objective scoring scheme. These tasks are open-ended tasks where participants are asked to produce a large number of different and original ideas from a basic target situation. For instance, respondents are asked to think and list all possible creative uses for a brick. Responses are typically scored for fluency (raw number of

ideas), flexibility (variety of ideas), and originality (often scored as statistical rarity in a given sample). Classically, it is admitted that divergent thinking is not tantamount of creativity; rather, it is an aptitude or a process ability that is at the core of creative potential (Runco, 2013). Divergent thinking tasks have also been criticized, even more severely, for being too reductionist, but empirical evidence shows that they nonetheless have some degree of real-life predictive validity (Plucker, 1999).

The classical scoring of divergent thinking tasks raises several problems (see Silvia et al., 2008 for a review). For instance, fluency, flexibility and originality indices are often excessively correlated, lacking discriminant validity. In order to overcome these limitations, Silvia et al. (2008) have proposed a scoring technique based on subjective ratings of ideas. This technique is called the Top 2 scoring method and goes as follows. First, participants complete divergent thinking tasks and select the two responses that they think are their most creative response. Then raters assess the creativity of responses using a 5-point scale and a definition of creativity that explicitly emphasizes the uncommonness, remoteness and cleverness of ideas. This method has proved to be valid and reliable with a relatively limited number of raters (2 or 3). It provides creativity scores that are less confounded with fluency than traditional originality scores (see also Benedek, Mühlmann, Jauk, & Neubauer, 2013).

Artistic and real-life creativity tasks. This class of creativity assessment has its origins in the psychosocial approach to creativity (Amabile, 1983). The basic idea here is to ask participants to actually complete a creative product – write a short story, make a drawing or a collage, etc. Typically, these products are assessed using the Consensual Assessment Technique (Amabile, 1982). The core idea of this method is to ask a panel of experts (typically artists) to rate the overall creativity of the products. Although effective, this method can be criticized for its use of a single item (i.e. “creativity”) and its lack of explicit definition; the whole procedure relies on experts’ implicit definition of creativity – a complete black box.

Moreover, this method is very costly. First, participants have to complete a real product, which can take a lot of time. Second, mobilizing several experts for the creativity ratings of hundreds of products can be very expensive and virtually impossible in large samples.

Alternative, complementary methods inspired by this line of research have been proposed. As regard the use of single "creativity" items, some researchers (e.g. Finke, Ward, & Smith, 1992) have recommended rating creativity indirectly through the assessment of quality and originality, both jointly contributing to creativity. As for the rater's selection, some research suggests that under certain circumstances, typically for everyday creativity, non-expert ratings can be used (Kaufman & Baer, 2012; Kaufman, Gentile, & Baer, 2005). Furthermore, combining the Top 2 scoring method, the use of peer-rating, and a variety of creativity tasks, Fürst and Lubart (2013; see also Fürst, in press) have provided evidence for the overall validity and reliably of this kind of approaches.

Insight tasks. Finally, insight tasks are also often used as a typical indicator of creativity. Unlike divergent thinking tasks, insight tasks have only one correct solution and generally involve an "aha" moment, when the solution appears suddenly, following a rather erratic and inscrutable search process. Typical examples of such tasks are the candle task or the nine dots problem (for an introduction on insight and detailed examples of typical tasks, see Sternberg & Davidson, 1996).

The Remote Associates Test (RAT; Mednick & Mednick, 1967) can also be seen as an insight task. It is based on verbal/conceptual association and consists of several items, each made of three stimulus words that are non-obviously, remotely related. The aim of the task is to find a fourth word that uncovers the underlying association between the three target words. This task, and the associative interpretation of creative thinking upon which it relies (Mednick, 1962) have been quite influential in creativity research. However, the RAT has also been criticized for being an indicator of (verbal) intelligence more than creativity (Lee,

Huggins, & Therriault, 2014). More generally, other studies have also questioned the relevance of insight tasks for creativity and documented their correlation with intelligence (Beaty, Nusbaum, & Silvia, 2014). In sum, this is a controversial issue and the debate about the relation between insight, intelligence, and creativity is not settled yet. Insight tasks such as the RAT would certainly benefit from further empirical exploration.

A multivariate scheme for the measurement of everyday creativity

Creativity has many facets and components and, accordingly, there is a very large number of instruments allowing the measurement of many types and forms of creativity. This surely is a good thing. However, this unstructured array of existing instruments does not match any integrative theory; there is no systematic measurement scheme that explicitly considers the person, process and product distinction. Yet, the multifaceted nature of creativity clearly calls for a multivariate, multi-instrument, multimethod operationalization. We now suggest elements of a solution to this issue, starting with a synthesis of the critical distinctions discussed at the beginning of the paper. We then propose a selection of instruments leading to a measurement scheme consistent with this synthesis.

Structuring the multiple facets of everyday creativity

It is obvious from the above that fundamentally, one can hardly speak of creativity "in general". It is essential to clarify what "type" or what "facet" of creativity is being talked about or measured. Table 1 offers a synthetic overview of these distinctions, focusing on those that are critical for this paper. This table presents, in columns, the *person-process-product* distinctions (slightly modified and enriched for our purposes). In rows, it presents the potential v. manifest distinction, the former being arguably domain-general (or at least with low domain specificity), while the latter is more domain-specific.

The first cell of Table 1 represents variables typical of *creative potential* that can be found in the *person* component. More specifically, this cell typically refers to personality

variables (e.g., Openness) and by extension breadth and depth of interest. Indeed, interest is probably a critical forerunner of creativity (Woo, Keith, Su, & Parrigon, 2017); interest leads to learning and domain-specific knowledge, which is essential for high-level creativity. The second cell of this "potential" row includes variables that represent creative *action*. This cell lists critical cognitive processes relevant to creativity, typically divergent thinking abilities and other facets of intelligence or convergent thinking. More generally, it also encompasses idea generation and selection abilities as a whole. The third cell of this row is empty (N.A.) because when dealing with potential creativity, the notion of actual creative products and achievements makes no real sense, though bits of creative ideas could be considered as creative products. However, for the sake of simplicity, this cell is left empty here[3].

[Insert Table 1 about here]

Similarly, the first cell of the second row (manifest creativity) is left empty as well, because it can be argued that personality variables only have little direct impact on manifest creativity. Rather, their effect is more likely to be indirect: personality essentially shapes overall creative potential which leads, along with other factors, to specific, actual activities and achievements. The second cell of this "manifest" row refers to such explicit, domain-specific creative *activity*, which can occur at different levels, from little-c to Big-C[4]. Finally, the *outcome* cell regroups both single observable creative products (whether created in a lab setting or in real life) and overall domain-specific achievements (measurable for instance with questionnaires assessing prizes received and critical or public acclaim).

Selecting and refining instruments

[3] On a related note, the notion of mini-c would fall somewhere between the action and outcome column. Indeed, mini-c can be seen as a process (i.e. "micro" generation or combination of ideas) or as a product ("embryonic" or "partial" creative output).

[4] One could argue that single creative ideas should also be listed in this cell, because they can also be manifest and thus observable. However, this cell is more specifically dedicated to significant, explicit, enduring and relatively complex creative activity, and not to small, occasional, isolated or accidental ones.

We now turn to the listing of potential instruments that could allow the assessment of the four cells just discussed. Ideally, several instruments should be used and mono-method approaches should be avoided (e.g. using tasks or questionnaires only). Indeed, questionnaire-based and task-based approaches have often been used separately, and it is quite rare to find systematic information on how the former are related to the latter (for a recent exception see Jauk, Benedek & Neubauer, 2014). Our multivariate and multi-method measurement scheme of everyday creativity proposes to use several instruments in order to assess the four key cells of Table 1:

- ***Potential creativity; Person Component***. First, as already mentioned, a measure of Openness/Intellect can probably be used as an overall acceptable proxy of creative personality. To avoid an excessively monolithic approach, a distinction between Openness and Intellect can be made. In addition, an inventory of the breadth and depth of creative interests could be added.
- ***Potential creativity; Process Component***. Classical divergent thinking tasks can be used to assess a critical process underlying creative potential. For simplicity, only one indicator (fluency) can be used, since it tends to correlate strongly with other indicators (flexibility and originality). Generation and Selection questionnaires can be used as additional indicators of the process side of creative potential.
- ***Manifest creativity; Activity Component***. As our review suggests, there are many questionnaires assessing creative activity. However, no existing questionnaire seems ideal; some do not provide explicit domain distinctions, others do not directly assess the intensity of activities. Below (Method section), we introduce a synthetic questionnaire that allows a precise assessment of creative activities – but also of interests and achievements – in several domains.

- ***Manifest creativity; Outcome Component***. A first way to assess creative outcome is to use lab tasks combining divergent thinking tasks and subjective ratings of creativity. Another way, complementary though controversial, is to add an insight task such as the RAT. Finally, questionnaires of creative achievement in several domains should be used as well.

On this basis, we now turn to the empirical part of the paper, introducing the instruments and their psychometric properties in more detail. We then provide a thorough analysis of the relations between them.

Method

Participants

The observations were collected across four compatible subsamples, resulting in a final total sample of 629 observations (72.4% women; mean age of 28.2 years, SD=8.58). Subsample 1 consists of 263 first-year undergraduate students at the University of Geneva (72.4% women; mean age of 24.2 years, SD=5.6). Subsample 2 consists of 120 persons from the general population (56% women; mean age of 42.8 years, SD=14.2); observations collected through an online procedure, using the Qualtrics panel & on-line sample service (https://www.qualtrics.com/online-sample). Subsample 3 is a convenience sample that consists of 59 persons from the general population (49.2% women; mean age of 31.6 years, SD=10.1), recruited at an international language fair held in Liège (Belgium). Subsample 4 consists of 187 first-year undergraduate students at the University of Geneva (68.3% women; mean age of 23.5 years, SD=8.7). In the overall sample, the majority of respondents were undergraduate students from the University of Geneva (71.8%). About half of the participants (subsamples 2 and 4; 48.9% of total) were remunerated for the time allocated to this study (≈5 USD and 25 USD, respectively).

Procedure

Observations were collected either with paper-and-pencil questionnaires (subsample 1; 42% of total) or with computerized questionnaires. The questionnaires, described below, had the same content in both cases. Additionally, 30% of the participants (subsample 4) completed creativity tasks, also described below. The total duration of data collection was approximately 30 minutes for the questionnaires, and 30 minutes for the tasks. This study was part of a larger-scale data gathering procedure that also included variables not discussed in this paper.

Material

***Questionnaires*.** Openness and Intellect were assessed with the scales developed by DeYoung et al. (2007). In their original versions, these are two 10-item scales; in the present study we used shorter versions, with 6 items for each scale (we selected the 2x6 items with the highest loading reported by DeYoung et al.).

Generation and Selection were measured with a short 12-item questionnaire (6 items for each factor, all scored positively) based on Fürst et al. (2012). Participants were asked to tell how frequently these behaviors applied to their work or leisure activities, using a scale from 1='almost never' to 5='very often.' Additional information about this scale is available in the Supplemental material.

To assess creative interests, activities and achievement, we synthetized many available instruments reviewed above. The objective of this synthesis was to design a questionnaire that (1) provides an explicit distinction between domains; (2) incorporates an explicit assessment of creative activities (based on time spent on these activities and seriousness about this practice); (3) offers an assessment of creative achievements that is more realistic than the one provided by the CAQ (which is far too demanding for the general population); (4) add the notion of creative interests, which arguably constitutes a critical precursor of activity and achievements.

Seven broad domains (or groups of domains) were retained: *music*, *literature & writing*, *performance art*, *visual arts*, *3D design*, *inventions & technical solutions*, and *science*. A more detailed description of these domains, along with the complete list of items, is available in the Supplemental material. For each of these domains, an initial filter yes/no question was asked in order to ascertain if participants had any *interest* in the area considered. If they answered 'yes', six additional questions related to the intensity of this interest were asked (e.g., "I like to learn new things in this domain"). Answers to these questions used a scale from 1='almost never' to 5='very often'.

In addition, for domains in which participants declared having an interest, another yes/no question was asked in order to clarify whether they had an *active practice* in a given domain. If a participant answered 'yes', two groups of five additional questions were asked, one referring to the intensity of practice of an activity, and another to his or her achievements in this domain. Sample items for activity are "I spend several hours every week practicing in this domain", "My practice in this domain is serious and important to me" (based on Verhaeghen et al., 2005). Answers to these questions were given using a scale from 1='almost never' to 5='very often'. Sample items for achievement are "I have won prizes or awards in this domain"; "I have been paid for my work in this domain" (based on Carson et al., 2005; Jauk et al., 2014). Answers were rated with the following scale: 1='never'; 1='once or twice'; 1='between 3 and 5 times'; 1='6 to 10 times'; 1='more than 11 times'.

Creativity tasks. Four creativity tasks were used: an insight task, a divergent thinking task, and two "real-life" creativity tasks (a writing and a drawing task). These tasks are described below, along with the computer interface we used for their administration.

Remote associates test. The Remote Associates Test (Mednick & Mednick, 1967) was included as an insight task. More specifically, we used the French 12-item version of this task developed by Maddux & Galinsky (2009). In this study, we used the first two items as

examples in order to introduce the task; hence final scores range from 0 to 10 only. As explained above, for each item, the principle of this task is to find a word that uncovers the remote relation between three target words (e.g. target: *blank-white-lines*; answer: *paper*). This task was presented on computer and time to complete it was limited to five minutes.

Boxes task. The divergent thinking task was the cardboard boxes task, adapted from Torrance (1966). In this task, we asked participants to "imagine a lot of original and clever things one could do with cardboard boxes". The instructions also emphasized that their ideas should be creative, different from the obvious and popular ones. Participants were also asked to find as many ideas as they could. This task was presented on computer and time was not limited – it was, however, suggested to spend about 5 minutes on it. As participants had written all their ideas and clicked a "next" button, a new screen was displayed, asking them to select what they thought were their two best ideas (i.e., procedure by Silvia et al. 2008). Participants could only tick boxes next to their ideas; they could not modify them at this point. (The rating procedure of the selected ideas is described further below.)

Writing task. The writing task was the "end of story" task by Lubart, Besançon, & Barbot (2011). In this task, participants are provided with the beginning of a story (i.e. "One afternoon, a mysterious character is walking in a park. Suddenly, he hears a funny noise"). They are then asked to imagine an ending for this story. As for the cardboard task, it was fully computerized and without time limitation. In contrast with the story task, however, participants were asked to produce only *one* ending for this story and the total length of the text was limited to 1,000 characters (a character counter was embedded in the interface). Instructions also emphasized creativity, and a few questions were provided to help participants to get started (e.g. "What is happening in the park? Who is this mysterious character? What will he find out?").

Drawing task. The last creativity task was the alien drawing task by Fink et al. (1992). In this task, participants were asked to draw an alien creature, with instructions emphasizing the creativity and originality of this creature, which should really differ from known animals and popular science-fiction creatures. The instructions also made clear that it was not important to know how to draw "properly"; participants were told that they could use short text explanations in order to clarify the role of specific features that could be difficult to draw. This task was administered on a Microsoft Surface tablet PC, which allowed participants to draw directly on the screen with a stylus. As in the writing task, participants were asked to create only one product (drawing), without any limitation of time, and a few questions were also provided in order to help participants getting started (e.g., In which environment does your creature live? What are its perception organs? In which way are they uncommon?).

Rating of creativity tasks. Aside from the RAT which was scored classically, the three other creativity tasks were scored using a new procedure. Basically, every participant rated the productions of previous participants, following a double-blind randomized procedure. This procedure, therefore, fully relies on peer-rating and allows a considerable reduction of the burden typical of classical scoring procedures. Indeed, it reduces it to practically zero. Participants only need to take a few additional minutes to rate the ideas randomly assigned to them for evaluation; since each participant only rates a few ideas, this amounts to only about five minutes. More specifically, each participant provided 24 ratings: three overall creativity ratings for two ideas (given by different participants) in the cardboard task; three ratings of end of story on three criteria (originality, quality, and overall creativity); and three ratings of drawings on the same three criteria. Persons invited to rate the products ("raters") were provided with detailed instructions regarding the overall rating process.

In the cardboard task, creativity was defined as the combination of *originality* (ideas or products that are rare, non-obvious, surprising, unexpected, remote) and *quality* (ideas or

products that are relevant, insightful, clever, ironic, funny). In the story task, the three criteria to be assessed were defined as follows. Originality: the end of the story feels really unique, e.g., something unexpected happens, events take a weird turn, the origin or consequences of the noise are special. Quality: the text is well-written and structured, words appear to be carefully chosen, one can sense a certain degree of elaboration, progression and coherence (and/or some kind of narrative break or clever twist in the end). Creativity was basically defined as the combination of originality and quality; however, participants were also told to trust their feelings and use their own conception of creativity. For the drawing task, the three criteria were similarly defined. Originality: the creature is really unique and extraordinary, it does not look like a human being or a typical science fiction creature from film. Quality: there is some cleverness in the conception of the creature: there is a notable degree of elaboration, not in a technical sense, but rather manifested through interesting or funny features, and/or components that interact with some features of the environment. Creativity, as in the story task, was defined as a combination of originality and quality, but participants were also given leeway to express their own impression about the creativity of each products.

Finally, raters were given products to be rated using an algorithm designed for the purposes of the study, which randomly assigns products (ideas, stories, or drawings) to raters. However, some constraints were embedded in the algorithm in order to ensure that: (1) raters never evaluated their own ideas; (2) three different persons rated a given product in a given task; and (3) the use of the same rater for one participant was kept to a minimum (for the majority of participants, the total of 24 scores were provided by six different raters or more [median = 6; first quartile = 5; third quartile = 7]). One intrinsic, but minor limitation of this system is that the last nine participants have partially or completely missing data. This, however, is a negligible problem, easily fixed by using a handful of additional raters at the end of the study.

Results

Creativity questionnaires

We start this section with an investigation of the creative interests, activities and achievements questionnaires' psychometric properties. We then investigate the relationships between these constructs and the other questionnaires variables[5].

Virtually all participants have at least a minimal interest in one domain; only 1.9% report no interest at all in any of the seven domains. The Cronbach's alpha across the seven domains for these six items systematically exceeds .80. Factor analysis performed on each domain shows that one factor systematically accounts for more than 50% of the variance. For activity and achievements, more participants did not report any of them (19.4%). As regards reliability, one exception aside, the Cronbach's alpha for all 14 scales (i.e., activity in seven domains and achievement in seven domains) all exceed .70. The only exception concerns the achievement scale in the 3D design (.62). Furthermore, exploratory factor analysis shows that the activity items and the achievement items almost systematically group together on distinct but correlated factors (median r = .38).

Let us now consider how these variables are related to the other questionnaire variables. This information is displayed in Table 2. One striking feature of this table is that all the variables are significantly correlated, with coefficients ranging from .23 (Intellect and creative achievement) to .88 (creative activity and creative achievement). The higher correlations are found between interest, activities and achievements, indicating that these constructs are, as expected, strongly interrelated. The four other variables are also quite highly correlated, especially Openness, Intellect and Generation. In addition, these four variables are consistently and positively related to interests, activities and achievements.

[5] Throughout this results section, we stay at a general level and do not investigate in detail the issue of domain specificity. Readers interested in this question can find additional analyses in the Supplemental material.

[Insert Table 2 about here]

Multiple regression shows that each of the four variables brings a unique, significant contribution when predicting creative interests. Openness is clearly the best predictor (β=.31; $p < .001$), followed by Selection (β=.17; $p < .001$), Generation (β=.13; $p = .004$) and – only marginally significant – Intellect (β=.08; $p = .053$). Together these variables predict 26% of the variance of creative interest. Analogous analyses in which the criterion is creative activities and achievements (regrouped in a single variable) reveal similar results, namely three positive and significant effects for Openness (β=.26; $p < .001$), Generation (β=.23; $p < .001$) and Selection (β=.09; $p = .02$), totaling 21% percent of explained variance.

Creativity tasks

We now turn to the results of the creativity tasks (Table 3), considering first the reliabilities (Cronbach's alpha) of the rating scores. For the Boxes task, the reliability of the two sets of three ratings is relatively low (.59 and .42), but the two total scores of each idea are nonetheless fairly well correlated ($r = .45$); total reliability, encompassing all 6 ratings, is .65. For the Story tasks, three sets of three ratings are available, one for originality, one for quality and one for overall creativity. As in the case of the Boxes task, the reliabilities of these sub-scores are quite low (about .42) but the three sub-scores are quite highly correlated (from .56 to .85); total reliability, encompassing all nine ratings, is .78. Finally, for the drawing tasks the reliability of the three sub-scores is slightly better than for the two other tasks (between .45 and .60), but still quite low. However, again, the sub-scores are highly correlated (from .64 to .75); total reliability, encompassing all nine ratings, is .82.

As regard correlation across tasks, all sub-scores of the boxes task are correlated with the sub-scores of the story task (r from .14 to .26) and the RAT (r from .20 to .30). The RAT is also marginally correlated with the originality scores of the drawing task ($r = .15$). The

fluency scores were not significantly correlated with the ratings, which is consistent with results by Silvia et al. (2008) and Benedek et al. (2013).

[Insert Table 3 about here]

Confirmatory factor analysis

In a last step of the analysis, we consider three structural equation models allowing a better understanding of the relationship between the creativity variables. We start with a measurement model of the creativity tasks, then we propose two integrative models encompassing most of the available variables.

Model 1 – Measurement model. The aim of the measurement model is to test to what extent it is realistic to regroup creativity tasks in a second-order factor of *Manifest creativity*. This model is based only on the Boxes and story Tasks. The drawing task is not included in this measurement model because this task was virtually unrelated to the two other tasks. Likewise, the RAT is not included either because of its different (and controversial) nature; here, we wanted a creativity factor based solely on ratings.

This model, depicted on Figure 1, relies on a logic similar to the one proposed by Silvia (2011). It includes four first-order factors: two factors extracted from the creativity ratings of the Top 2 ideas in the Boxes task (Bi1 and Bi2), and two factors extracted from the rating in the Story tasks, namely ratings of quality (Sq) and rating of originality/creativity (So/c). The originality and creativity ratings, initially distinct but very strongly correlated (r = .85), were averaged in a single indicator. Several residual correlations are set free, namely, between the three pairs of ratings provided by the same group of raters in the Boxes task (i.e., the same group of raters has evaluated idea 1 and idea 2); and between the three pairs of ratings provided by the same group of raters in the Story task (i.e., the same group of raters has evaluated the quality and originality/creativity of the story). Finally, the second factor of Manifest creativity loads on the four first-order factors just described.

The fit of this model is very good ($\chi^2_{(44)}$ = 53.4, p = .157; χ^2/DF = 1.21; RMSEA = .036, 95% CI [0, .066]; SRMR = .053; CFI = .975)[6]. Unsurprisingly, the loadings on the first-order factors are modest (but all significant) and the residual correlations, reflecting "rater effects" are substantial. However, the loadings on the second-order factor are all quite high. Remarkably, the reliability of this factor, as estimated with the *H* coefficient (Hancock & Mueller, 2001) is .83, which is good to excellent. Overall, this confirms the fact that it is possible to extract a meaningful, reliable factor of manifest creativity based on these two tasks, even though the number of raters per task is relatively low.

[Insert Figure 1 about here]

Model 2 – Integrative model. Based on the foregoing analysis, the next step was to specify and test a model integrating the creativity variable based on questionnaires. The original specification of this model was based on Table 1. However, an initial round of testing showed that latent variables representing creative potential (i.e., the two cells in the first line of Table 1) were very highly correlated (r = .90; p < .001); hence it was more parsimonious to represent creative potential with a single factor. In addition, this initial round of testing also showed that creative activities and creative achievements were best grouped together (i.e. self-reported manifest creativity), and that creative interests loaded both on creative potential and self-reported manifest creativity. The final model represented on Figure 2 has a fairly good fit ($\chi^2_{(157)}$ = 221.9, p = .0005; χ^2/DF = 1.41; RMSEA = .026, 95% CI [.017, .033]; SRMR = .062; CFI = .973) and represents the comprehensive assessment of creativity we aimed for – one factor of creative potential and two distinct factors of manifest creativity, all positively correlated (although relatively modestly).

[6] An alternative measurement model, with two intermediary second-order factors (representing the Boxes and Story tasks) and manifest creativity represented by a third-order factor (loading on the two tasks) fit the data slightly better ($\chi^2_{(45)}$ = 51.43, compared to $\chi^2_{(44)}$ = 53.4 for the model depicted on Figure 1). However, this model was deemed needlessly complicated given the negligible difference of fit.

[Insert Figure 2 about here]

Model 3 – Alternative model. Finally, an alternative, complementary model was estimated. The aim of this model was to better inform the fairly complex relations between all the variables at hand. The central idea here was to "deconstruct" the self-reported manifest creativity factor and to consider creative interests, activities and achievement separately – investigating their relations together, and with the two other key factors identified above (i.e., creative potential and manifest creativity). This model, represented in Figure 3, has a good fit, similar to the previous ones ($\chi^2_{(176)} = 223.01$, $p = .0026$; $\chi^2/DF = 1.32$; RMSEA = .043, 95% CI [.026, .057]; SRMR = .062; CFI = .941).

[Insert Figure 3 about here]

Specifically, this model shows that creative potential is a strong predictor of creative interests ($\beta = .57$) which are, in turn, a strong predictor of creative activities ($\beta = .50$). Creative activities are very closely related to creative achievements ($\beta = .88$) which are themselves a positive predictor of the task-based manifest creativity factor ($\beta = .27$). However, surprisingly, the creative potential factor is an even better predictor of the performance in creativity tasks; when included in the model, this path ($\beta = .28$) virtually nullifies the .27 prediction of creative achievement. Finally, the RAT is also a non-negligible predictor of manifest creativity in tasks.

Discussion

The aim of this paper was to propose a general method of creativity measurement that is strongly multivariate, theoretically driven, and realistic. In what follows, we critically review the results and discuss them in light of theoretical elements provided in the introduction. We also provide several avenues for future research, from possible ameliorations to various adaptations and extensions.

Several questionnaires were used in this study, and all of them showed good to excellent psychometric properties. Reliability figures were high for all scales, and exploratory factor analysis showed that all instruments roughly had the expected factorial structure. Remarkably, all the scores' distributions – including mean achievement across all domains – were normal or almost normal. This is a direct result of our willingness to avoid an overrepresentation of items focused on very high achievement (e.g., national and international recognition). As expected, the relations between all these questionnaire variables were numerous, all positive and significant.

First, as suggested by Woo et al. (2017), creative interests were found to be highly correlated with creative activities and achievements. Consistent with the notion that interests are an important aspect of creative potential, this scale was also positively correlated with Openness, Intellect, Generation and Selection. As shown by multiple regression analysis, each of these four variables contributes uniquely to the prediction of creative interests. Activities and achievements were also strongly correlated with Openness, Intellect, Generation and Selection. Moreover, again as shown by multiple regression, almost all these variables contribute uniquely to the prediction of activities and achievements. Along with results involving creative interests, this provides evidence of the discriminant validity of all these instruments; each one captures a different and complementary aspect of creativity. (Analysis by domains, although not detailed in this paper, also shows that all four variables play an important role in predicting creativity in specific domains.)

Several creativity tasks were also included in his study, namely one divergent thinking task (Boxes), two real-life like creativity tasks (Story and Drawing) and the Remote Associates Test (RAT). Several indices were derived from these tasks. The Boxes task led to a classical fluency score and to several Top 2 idea ratings. The Story and Drawing tasks led to

several ratings of originality, quality and overall creativity. All these ratings were obtained using a relatively new subjective scoring method, based on peer evaluations only.

As regards reliability, this method has led to mixed but encouraging results. Sub-scores in each task had relatively poor reliability, in the .40-.60 range. However, the reliability of the overall score in each task (based on six or nine ratings in total) was acceptable to good (from .65 to .82). We might concur here with Benedek et al. (2013) that more ratings should perhaps be used in order to achieve higher reliabilities (either using four raters instead of three and/or a "Top 3 ideas" paradigm instead of a "Top 2"). However, despite the relatively limited reliability of each task's sub-scores, the measurement model based on the Boxes and Story tasks (Figure 1) has made it clear that it was definitely possible to construct a latent variable of manifest creativity with high reliability (over .80). Hence it appears that adequate modelling can considerably enhance measurement quality.

Ultimately, the core goal of this paper was to propose a measurement scheme consistent with the synthesis provided in Table 1. Although this synthesis has led us to identify four groups of variables, while confirmatory factor analysis suggests the existence of three distinguishable factors (i.e., Figure 2), the results fit well in the theoretical framework. The first factor (Creative Potential) encompasses variables listed in the first row of Table 1, namely Openness, Intellect, creative interests, Generation, Selection, and fluency. The second factor (self-reported Manifest Creativity) is represented by creative activities and achievements (two cells in the second row). The third factor (task-based Manifest Creativity) represents the Boxes and Story tasks, in line with the description of the fourth cell of Table 1 (Manifest creativity; Outcome Component).

The final model (Figure 3) provides interesting additions and clarifications. This relatively complex mediation model shows that creative interests are a key mediator between Creative Potential and creative activities. Hence the consideration and measurement of

creative interests can be seen as very valuable. This model also shows that the Creative Potential latent variable has strong predictive power – actually stronger than the impact of creative achievement. This clearly shows the importance of the notion of creative potential. In sum, this third model provides a comprehensive and theoretically meaningful model of creativity. Despite its apparent complexity, it is relatively parsimonious (only a handful of key relations synthetizes the complex manifold positive correlations between several variables).

In this general picture, two tasks had a somewhat special status: the RAT and the Drawing task. The RAT was not correlated to any task other than the Story. Given recent studies on this task and its relation to creativity (Beaty et al., 2014; Lee et al., 2014) null results were in fact to be expected. In this context, the correlation we found actually pleads in favor of the validity of the RAT. Moreover, the RAT did predict the tasked-based Manifest Creativity factor in the last model (Figure 3). The drawing task, although it was the task with the higher reliability, was not correlated to any other total score, arguably because of domain specificity (most people find drawing difficult)[7].

Overall, despite these encouraging results, the road towards an ideal multivariate scheme of creativity assessment is still long. However, the field of creativity measurement has, over the past ten years, been revitalized by solid and innovative work, contributing to the development of both tasks and questionnaires. Below, we assess how the present contribution relates to this research and identify priorities for future research.

Our interests, activities and achievements questionnaire, offering a unified approach of these three themes, is a practical and flexible instrument that synthesizes various lines of work (Jauk et al., 2014; Silvia et al., 2012; Verhaeghen et al., 2005). The progression from interests to achievements, using two yes/no filter questions and three five-item scales, provides rich

[7] It is worth noting that in a "domain-specific" version of the model represented in Figure 2, we found that self-reported manifest creativity in the visual arts did predict the creative score in the Drawing task (β = .20, p = .03), thus suggesting that this Drawing task is valid, but has very high domain specificity.

data and is also practical for participants – they only answer items that are relevant to them, thus saving time and effort. As regards the question of domains, this scale, as introduced here, is mainly focused on art, as well as on technique and science. However, other domains could easily be added (e.g. cooking or business). Finer-grain domain distinctions could also be made. For instance, the visual arts domain could be split up in drawing, painting, photography, graphical design, etc. The fact that items are almost identical for all domains makes such variations very easy, requiring only minimal adjustments. Future studies can explore these options and provide mutually compatible replications and extensions.

As regards tasks, our approach combines strands of research on the use of non-expert or peer ratings and alternative scoring of divergent thinking tasks (Baer, Kaufman, & Gentile, 2004; Benedek et al., 2013; Fürst & Lubart, 2013; Fürst, in press; Kaufman & Baer, 2012; Kaufman, Baer, Cropley, Reiter-Palmon, & Sinnett, 2013; Silvia et al., 2008). This literature suggests that peer-rating is acceptable for everyday creativity (especially for the evaluation of short texts and the rating of divergent thinking tasks) and if the raters are provided with sufficiently precise instructions. Our approach follows and expands on these contributions. It also provides a number of practical advantages. First, given that participants at time *t* act as evaluators of participants at *t-1*, the ratings are produced almost instantly, as the study progresses. There is no need to process ideas for evaluation, no need to regroup identical or similar ideas, no need to organize complex sessions of idea ratings. Moreover, this method relies on dozens (or hundreds) of raters; the number of raters grows as the study grows, hence putting no limitation whatsoever on sample size. This system would work equally well in a very large study (e.g., n=1000 or more) which would otherwise be a logistic nightmare. The randomized assignment of products to raters, combined with constraints that ensure that each participant is rated by several other participants, guarantees that the creativity scores rely on a substantial variety of people. This ensures the diversity and robustness of the ratings.

On this basis, along with the use of other existing instruments, a wide range of customized multivariate schemes of creativity assessment can be envisaged. In particular, the person (or personality) component of creative potential can be assessed with questionnaires of the type proposed here as well as with analogous/complementary scales (e.g., a short Big Five questionnaire, allowing the assessment of Plasticity and Stability). The process (or cognitive) component of Creative Potential should also be assessed with classical fluency scores along with intelligence tasks. Indeed, it seems reasonable to consider intelligence as a component of creative potential (see, e.g. Nusbaum & Silvia, 2011; Jauk et al., 2014). Measurement schemes should also make a distinction between domains and levels of creativity. At the everyday level, manifest creativity could be assessed with questionnaires and tasks such as those introduced here (for complements or alternatives, see for instance Benedek et al., 2013; Silvia, Martin, & Nusbaum, 2009; Silvia et al. 2012; Jauk et al. 2014). For higher levels such as Big-C, questionnaires like the CAQ (Carson et al., 2005) and the consensual assessment-based expert ratings (Amabile, 1982) remain the gold standard.

Conclusion

The main thrust of this paper was to provide a multivariate and multimethod assessment of everyday creativity, synthetizing several research traditions and including personality, cognitive styles, creative interests, activities and achievements, as well as several creativity tasks encompassing several domains and scoring methods. This results in an empirical design that fits into a one-hour protocol which avoids the tedious and costly processing typically associated with the subjective rating of creative products. Moreover, this design relies on a general framework that builds on most of the theoretical inputs of past research in the field. Our hope is that the approach proposed here can make a useful contribution to the development of multivariate and theoretically driven perspectives on creativity assessment.

References

Amabile, T. M. (1983). The Social Psychology of Creativity: A Componential Conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357-376.

Baer, J. (1993). *Creativity and divergent thinking. A task-specific approach.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the Consensual Assessment Technique to Nonparallel Creative Products. *Creativity Research Journal*, *16*(1), 113-117.

Batey, M. (2007). *A psychometric investigation of everyday creativity* (Unpublished doctoral dissertation). University College, London.

Batey, M., & Furnham, A. (2006). Creativity, Intelligence, and Personality: A Critical Review of the Scattered Literature. *Genetic, Social, and General Psychology Monographs*, *132*(4), 355-429.

Beaty, R. E., Nusbaum, E. C., & Silvia, P. J. (2014). Does insight problem solving predict real-world creativity? *Psychology of Aesthetics, Creativity, and the Arts*, *8*(3), 287-292.

Beghetto, R. A. (2015). "Learning as a creative act", in T. Kettler (Ed.), *Modern Curriculum for Gifted and Advanced Learners,* p. 111-127. New York: Routledge.

Beghetto, R., & Kaufman, J. C. (2007). Toward a Broader Conception of Creativity: A Case for "mini-c" Creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *1*(2), 73-79.

Benedek, M., Mühlmann, C., Jauk, E., & Neubauer, A. C. (2013). Assessment of Divergent Thinking by means of the Subjective Top-Scoring Method: Effects of the Number of Top-Ideas and Time-on-Task on Reliability and Validity. *Psychology of aesthetics, creativity, and the arts*, *7*(4), 341-349.

Brophy, D. R. (2000). Comparing the Attributes, Activities, and Performance of Divergent, Convergent, and Combination Thinkers. *Creativity Research Journal*, *13*(3-4), 439-455.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, Validity, and Factor Structure of the Creative Achievement Questionnaire. *Creativity Research Journal*, *17*(1), 37-50.

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality*, *46*(1), 63-78.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between Facets and Domains: 10 Aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880-896.

DeYoung, C. G., Shamosh, N. A., Green, A. E., Braver, T. S., & Gray, J. R. (2009). Intellect as distinct from openness: Differences revealed by fMRI of working memory. *Journal of Personality and Social Psychology*, *97*(5), 883-892.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin*, *136*(5), 822-848.

Dollinger, S. J. (2003). Need for uniqueness, need for cognition and creativity. *The Journal of Creative Behavior*, *37*(2), 99-116.

Eysenck, H. J. (1995). *Genius: The Natural History of Creativity*. New York: Cambridge University Press.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA, US: MIT Press.

Fürst (in press). Measuring creativity with planned missing data. *Journal of Creative Behavior*.

Fürst, G., Ghisletta, P., & Lubart, T. (2012). The creative process in visual art: A longitudinal multivariate study. *Creativity Research Journal*, *24*(4), 283-295.

Fürst, G., Ghisletta, P., & Lubart, T. (2016). Toward an Integrative Model of Creativity and Personality: Theoretical Suggestions and Preliminary Empirical Testing. *The Journal of Creative Behavior*, *50*(2), 87-108.

Fürst, G., Ghisletta, P., & Lubart, T. (2017). An experimental study of the creative process in writing. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(2), 202‑215.

Fürst, G., & Lubart, T. (2017). An Integrative Approach to the Creative Personality: Beyond the Big Five Paradigm. In G. J. Feist, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *The Cambridge Handbook of Creativity and Personality Research* (1 edition, p. 140‑164). New York, NY: Cambridge University Press.

Fürst, G., & Lubart, T. I. (2013). *The measurement of creativity : Consensual assess- ment technique and divergent thinking tasks.* Paper presented to the First International Conference on the Science of Creative Thinking, Bologna, Italy, September 29-October 1.

Gajda, A., Beghetto, R. A., & Karwowski, M. (2017). Exploring creative learning in the classroom: A multi-method approach. *Thinking Skills and Creativity*, 24, 250-267.

Glăveanu, V. P. (2013). Rewriting the language of creativity: The Five A's framework. *Review of General Psychology*, 17(1), 69-81.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology*, *37*(8), 1398‑1405.

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*, 444‑54.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*(4), 267‑293.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), Structural equation modeling: Present and future (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Hocevar, D. (1979). *The development of the Creative Behavior Inventory (CBI)*. Paper presented at the annual meeting of the Rocky Mountain Psychological Association.

Jauk, E., Benedek, M., & Neubauer, A. C. (2014). The Road to Creative Achievement: A Latent Variable Model of Ability and Personality Predictors. *European Journal of Personality*, *28*(1), 95-105.

Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 298-308.

Kaufman, J. C., & Baer, J. (2005). The Amusement Park Theory of Creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse.* (p. 321-328). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Kaufman, J. C., Baer, J., Cropley, D. H., Reiter-Palmon, R., & Sinnett, S. (2013). Furious activity vs. understanding: How much expertise is needed to evaluate creative work? Psychology of Aesthetics, Creativity, and the Arts, 7, 332-340.

Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal*, *24*(1), 83-91.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*(1), 1-12.

Kaufman, J. C., Beghetto, R. A., Baer, J., & Ivcevic, Z. (2010). Creativity polymathy: What Benjamin Franklin can teach your kindergartener. Learning & Individual Differences, 20, 380-387.

Kaufman, J. C., Cole, J. C., & Baer, J. (2009). The Construct of Creativity: Structural Model for Self-Reported Creativity Ratings. *Journal of Creative Behavior*, *43*(2), 119-132.

Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, *61*(5), 622-629.

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, *133*(3), 464-481.

Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence? Examining internal and external structure validity evidence of the Remote Associates Test. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(4), 446-460.

Lubart, T. I., Besançon, M., & Barbot, B. (2011). *Evaluation du Potentiel Créatif [Evaluation of Creative Potential]*. Paris, France: Hogrefe.

Maddux, W. W., & Galinsky, A. D. (2009). Cultural borders and mental barriers: The relationship between living abroad and creativity. *Journal of Personality and Social Psychology*, *96*(5),

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, *52*, 1258-1265.

McKay, A. S., Karwowski, M., & Kaufman, J. C. (2017). Measuring the muses: Validating the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts*, *11*(2), 216-230.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *3*, 220-232.

Mednick, S. A., & Mednick, M. T. (1967). *Examiner's manual, Remote Associates Test*. Boston, MA: Houghton Mifflin.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? *Intelligence*, *39*(1), 36-45.

Oleynick, V. C., DeYoung, C. G., Hyde, E., Kaufman, S. B., Beaty, R. E., & Silvia, P. J. (2017). Openness/Intellect: The Core of the Creative Personality. In G. J. Feist, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *The Cambridge Handbook of Creativity and Personality Research* (1 edition, p. 9-27). New York, NY: Cambridge University Press.

Perrine, N. E., & Brodersen, R. M. (2005). Artistic and Scientific Creative Behavior: Openness and the Mediating Role of Interests. *The Journal of Creative Behavior*, *39*(4), 217-236.

Plucker, J. A. (1999). Is the Proof in the Pudding? Reanalyses of Torrance's (1958 to Present) Longitudinal Data. *Creativity Research Journal*, *12*(2), 103.

Plucker, J. A., & Beghetto, R. A. (2004). Why creativity is domain general, why it looks domain specific and why the distinction does not matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From Potential to Realization* (p. 153-168). Washington DC: American Psychological Association.

Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, *55*, 657-687.

Runco, M. A. (2013). *Divergent Thinking and Creative Potential*. New York: Hampton Press.

Schweizer, T. S. (2006). The Psychology of Novelty-Seeking, Creativity and Innovation: Neurocognitive Aspects Within a Work-Psychological Perspective. *Creativity and Innovation Management*, *15*(2), 164-172.

Rhodes, M. (1961). An analysis of creativity. Phi Delta Kappan, 42, 305–311.

Silvia, P. J. (2011). Subjective scoring of divergent thinking: Examining the reliability of unusual uses, instances, and consequences tasks. Thinking Skills and Creativity, 6(1), 24-30.

Silvia, P. J., Martin, C., & Nusbaum, E. C. (2009). A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking. *Thinking Skills and Creativity*, *4*(2), 79-85.

Silvia, P. J., Wigert, B., Reiter-Palmon, R., & Kaufman, J. C. (2012). Assessing creativity with self-report scales: A review and empirical evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 19-34.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., … Richard, C. A. (2008). Assessing Creativity With Divergent Thinking Tasks: Exploring the Reliability and Validity of New Subjective Scoring Methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 68-85.

Sternberg, R. J. (1997). *Thinking styles*. New York: Cambridge University Press.

Sternberg, R. J., & Davidson, J. E. (Eds.). (1996). *The Nature of Insight*. Cambridge, Mass.: The MIT Press.

Sternberg, R. J., Grigorenko, E. L., & Singer, J. L. (Eds.). (2004). *Creativity: From Potential to Realization*. Washington, DC: American Psychological Association.

Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking-Norms-Technical Manual Research Edition-Verbal Tests, Forms A and B-Figural Tests, Forms A and B*. Princeton, NJ: Personnel Press.

Verhaeghen, P., Joorman, J., & Khan, R. (2005). Why we sing the blues: the relation between self-reflective rumination, mood, and creativity. *Emotion*, *5*(2), 226-232.

Woo, S. E., Keith, M., Su, R., & Parrigon, P. (2017). The curious dynamic between openness and interests in creativity. In G. J. Feist, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *The Cambridge Handbook of Creativity and Personality Research* (1 edition, p. 44-63). New York, NY: Cambridge University Press.

Tables

Table 1. Theoretical overview of the measurement scheme

	Person Personality, Interests	**Action** Processes, Activities	**Outcome** Products, Achievements
Potential creativity *(low domain specificity)*	Personality traits (chiefly Openness) Depth and breadth of interests.	Divergent thinking (fluency), intelligence. Idea Generation and Selection abilities.	N.A.
Manifest creativity *(high domain specificity)*	N.A.	Explicit and enduring creative activity, whether at a *little-c*, *Pro-C* or *Big-C* level.	Performance in lab creativity tasks. *Pro-C* or *Big-C* real-life creative achievements.

Table 2. Descriptive statistics and correlations across questionnaires

	Openness	Intellect	Generation	Selection	C. Interest	C. Activity	C. Achiev.
Descr. Stats.							
Mean	3.65	3.56	3.40	3.56	1.79	0.80	0.41
Std. dev.	0.64	0.72	0.75	0.72	0.89	0.65	0.36
Skewness	-0.29	-0.30	-0.19	-0.20	0.26	0.69	1.16
Kurtosis	-0.17	-0.21	-0.06	0.17	-0.42	-0.13	1.57
Reliability	.70	.73	.87	.85	.81-.87	.72-.84	.62-.90
Correlations							
Openness	1						
Intellect	.42***	1					
Generation	.47***	.51***	1				
Selection	.29***	.26***	.37***	1			
C. Interests	.45***	.32***	.37***	.33***	1		
C. Activities	.41***	.25***	.38***	.25***	.63***	1	
C. Achiev.	.34***	.23***	.37***	.24***	.54***	.88***	1

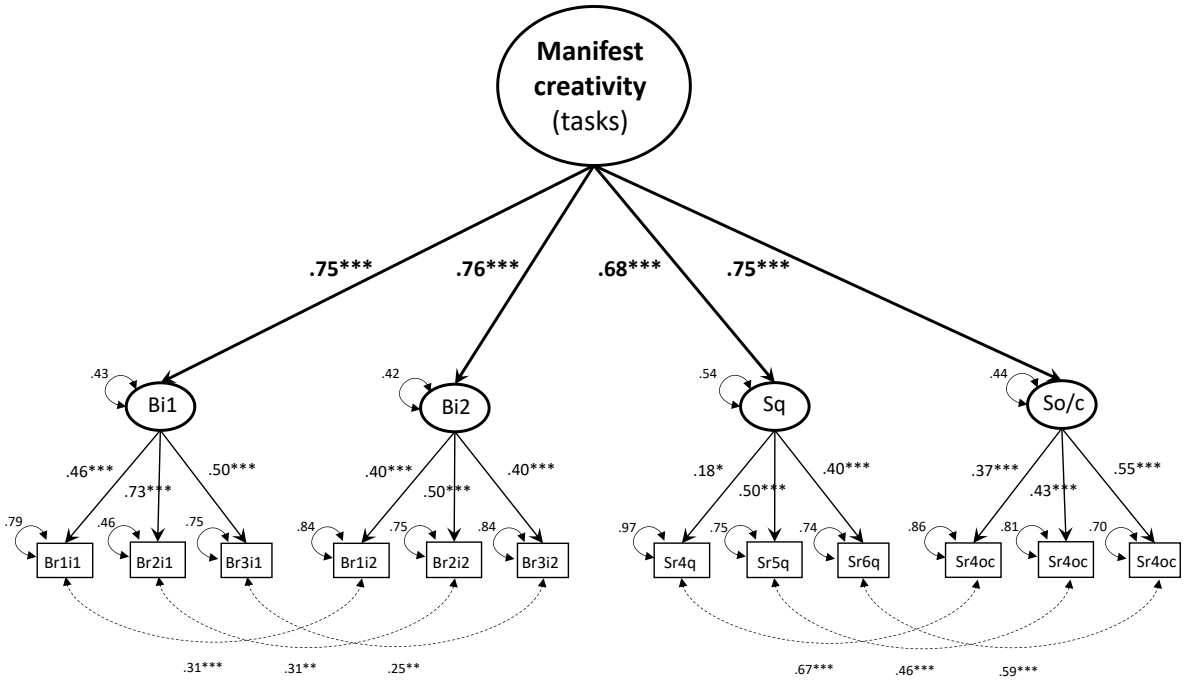
Note. ***: $p<.001$.

Table 3. Descriptive statistics and correlations across the creativity tasks

	RAT	Boxes			Story			Drawing		
		fluency	idea 1	idea 2	org.	qual.	crea.	org.	qual.	crea.
Descr. Stats.										
Mean	4.24	4.83	2.77	2.96	3.20	3.09	3.16	3.02	2.91	3.02
S.D.	2.37	2.27	0.91	0.82	0.83	0.78	0.78	0.86	0.80	0.76
Skewness	.24	1.40	0.04	-0.08	-0.19	-0.06	-0.20	-0.11	-0.10	-0.40
Kurtosis	-.63	2.62	-0.62	-0.50	-0.50	-0.58	-0.23	-0.46	-0.42	-0.28
Reliability	.74	-	0.59	0.42	0.43	0.42	0.43	0.45	0.60	0.45
Correlations										
Boxes	1									
fluency	.04	1								
idea 1	.04	-.03	1							
idea 2	.10	.07	.45***	1						
Story										
originality	.20*	.04	.26***	.20*	1					
quality	.30***	.04	.20*	.14†	.56***	1				
creativity	.22**	.03	.22**	.15†	.85***	.58***	1			
Drawing										
originality	.15†	.04	.01	.11	.06	-.01	.02	1		
quality	.03	-.02	-.07	.06	.10	.08	.08	.64***	1	
creativity	.11	-.01	.03	.11	.09	.01	.07	.85***	.75***	1

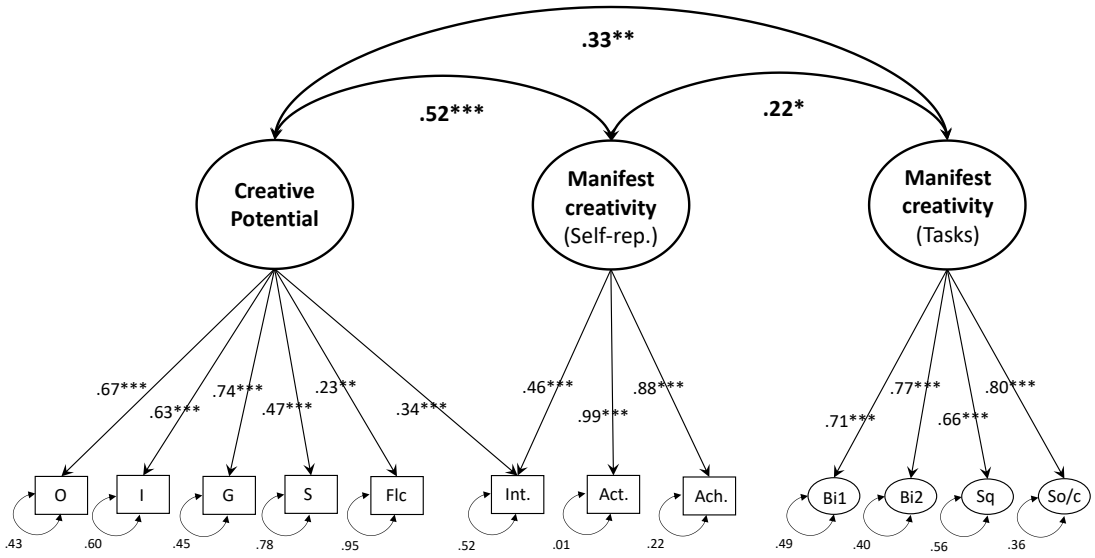
Note. †: $p<.10$*: $p<.05$; **: $p<.01$; ***: $p<.001$.

Figure 1. Measurement model of the three creativity tasks



Note. **B** = Boxes task; **i1** = creativity of first idea; **i2** = creativity of second idea; **S** = Story task; **q** = quality of text; **o/c** = originality and creativity of text (averaged in a single indicator); **r1**, **r2**, etc. = rater 1, rater 2, etc. *: $p<.05$; **: $p<.01$; ***: $p<.001$. See text for further information and model fit.

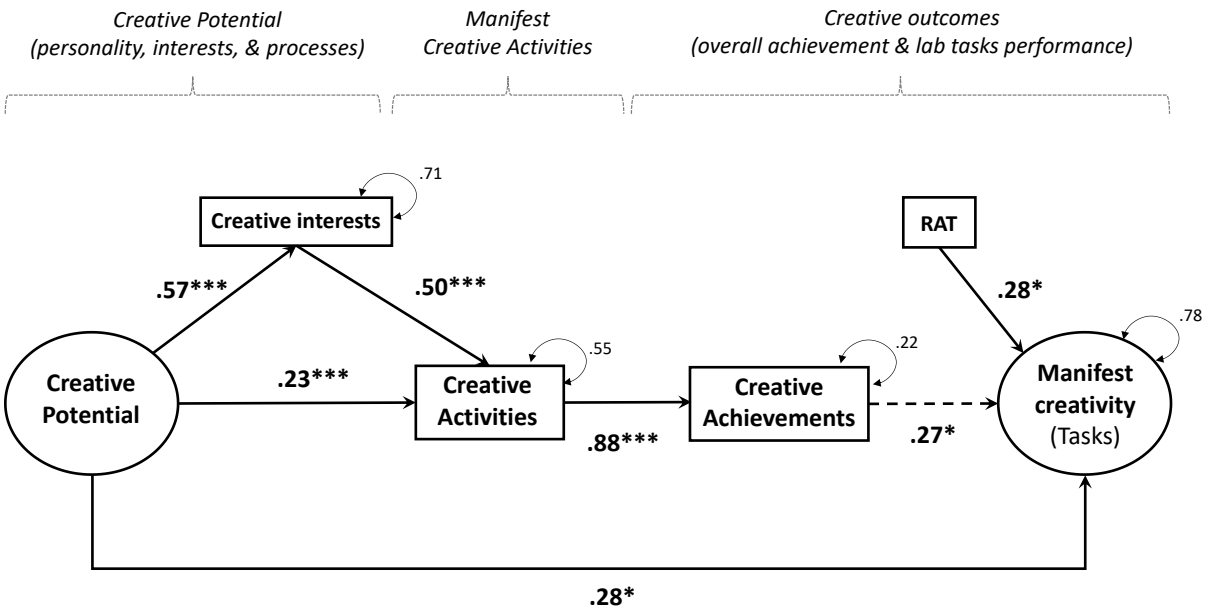
Figure 2. Integrative model



Note. **O** = Openness; **I** = Intellect; **G** = Generation; **S** = Selection; **Flc** = Fluency; **Int.** = creative interests; **Act.** = creative activities; **Ach.** = creative achievements; **Bi1** = Boxes task, creativity of idea 1; **Bi2** = Boxes, creativity

of idea 2; **Sq** = Story task, quality; **So/c** = Story task, originality/creativity. The Bi1, Bi2, Sq and So/c factors are specified as in Fig. 1. *: p<.05; **: p<.01; ***: p<.001. See text for further information and model fit.

Figure 3. Alternative model



Note. The *Creative Potential* latent variable is specified as in Fig. 2, but without creative interest as an indicator; the *Manifest Creativity* latent variable is specified as in Figure 2. *: p<.05; ***: p<.001. See text for model fit.