Psychology of Aesthetics, Creativity, and the Arts

The Role of Asking More Complex Questions in Creative Thinking

Tuval Raz, Roni Reiter-Palmon, and Yoed N. Kenett

Online First Publication, December 7, 2023. https://dx.doi.org/10.1037/aca0000658

AMERICAN PSYCHOLOGICAL ASSOCIATION
DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

ISSN: 1931-3896
https://doi.org/10.1037/aca0000658

The Role of Asking More Complex Questions in Creative Thinking

Tuval Raz[1], Roni Reiter-Palmon[2], and Yoed N. Kenett[1]
[1] Faculty of Data and Decision Sciences, Technion—Israel Institute of Technology
[2] Department of Psychology, University of Nebraska

Question asking has been a critical tool for teaching and learning since the time of Socrates and is important in the creative problem-solving process. Yet, its role in creativity has insofar not been thoroughly explored. The current study assessed the role of question asking in the creative process. A correlational preregistered design was used to administer the alternative questions task (AQT) to explore its relation to cognitive and creative divergent thinking tasks. In the AQT—which is based on Torrance's unusual questions task—participants are asked to generate creative and unusual questions for common objects. Responses are rated for their question level using the Bloom's taxonomy, a widely accepted guideline in designing examination questions of differing levels of complexity, as well as their subjective and objective creativity. A significant positive relation between AQT question level and objective and subjective creativity scores was found: Higher, more complex questions were more creative, with the inverse effect for lower-level questions. We interpret these findings as supporting the hypothesis that higher question complexity is related and predictive of creative ability. A second study replicated and generalized our findings. Thus, our findings uniquely highlight the role of question asking, and especially question complexity, in creativity.

Keywords: creativity, question asking, Bloom taxonomy, divergent thinking

Creativity entails both idea originality and appropriateness (Runco & Jaeger, 2012). Thus, creative ideas or solutions require skilled problem solvers to search their memory and "move away" from common ideas toward ideas that are more novel or conceptually distant (Abraham & Bubic, 2015; Beaty & Kenett, 2023; Benedek et al., 2023; Kenett, 2018; Kenett & Faust, 2019; Volle, 2018). However, creative thinking is also critically motivated by information-seeking behaviors that are driven by curiosity and the personality trait openness to experience (Kenett et al., 2023). Such information-seeking tendencies likely promote problem finding, the first stage in the creative problem-solving process (Reiter-Palmon & Robinson, 2009).

Problem finding is considered the first stage of the creative problem-solving process but is still far from being understood (Okuda et al., 1991). It can be defined in general terms as the process or processes that precede problem solving (i.e., occurring before a problem can be solved). Operationally, it may involve the identification of a problem or the definition of an ambiguous situation into a workable problem or the raising of questions from ill-defined problem situations (Getzels, 1979; Runco & Nemiro, 1994).

Editor's Note. Roni Reiter-Palmon is the book editor for the journal.

Roni Reiter-Palmon https://orcid.org/0000-0001-8259-4516
Yoed N. Kenett https://orcid.org/0000-0003-3872-7689

This work was partially supported by the U.S.–Israel Binational Science Fund (Grant 2021040). The authors declare no conflicts of interests. The authors thank Edgar Onea for early discussions on this project.

Correspondence concerning this article should be addressed to Yoed N. Kenett, Faculty of Data and Decision Sciences, Technion—Israel Institute of Technology, Kiryat Hatechnion, Haifa 3200003, Israel. Email: yoedk@technion.ac.il

Ill-defined problems often entail multiple, even sometimes conflicting, goals (Getzels, 1979; Schraw et al., 1995). There are multiple possible approaches to solve ill-defined problems. Before ideas can be generated, then evaluated, and selected for implementation, a process is needed to conceptualize and structure the ill-defined problem. During the problem finding process, an individual identifies, assesses, and structures a problem (Reiter-Palmon & Robinson, 2009). Constructing a new or unique approach to solving a problem makes the generation of creative ideas for solving the problem possible (Yang et al., 2022).

Past research indicates that problem finding and construction are positively related to creative problem solving (Mumford et al., 1991, 1994; Reiter-Palmon et al., 1997, 1998) and to divergent thinking measures of creativity (Abdulla et al., 2020; Alabbasi et al., 2023; Arreola & Reiter-Palmon, 2016). Reiter-Palmon et al. (1997, 1998) have found that people who excel at problem-finding tend to restate problems as questions, resulting in ambiguous or ill-defined problems. The researchers measured problem-finding ability based on the quality and originality of these restatements. However, much is still unknown about the specific types of these questions asked and their relationship to creative thinking.

The aim of the current study was to reintroduce the assessment of question asking in creativity research by using a creative questions task, and by utilizing current computational semantic distance methods to quantitatively assess participants' questions and their creativity. Critically, we focus on the role of question complexity in creativity, via an established taxonomy of learning objectives (the Bloom taxonomy).

Question Asking

An important but understudied part of creativity that likely facilitates information seeking behavior is question asking ability. In fact, question asking has been shown in the past to be part of the creative problem-solving process in children (Torrance, 1970). One of the tasks historically used to assess question asking was the unusual questions task (UQT), based on Burkhart's (1961) divergent questions task, the UQT was part of the original version of the classic creativity assessment battery, the Torrance Test of Creative Thinking (TTCT; Torrance, 1966). The TTCT is one of the most widely used and studied measures of creative idea generation (Grajzel et al., 2022). The UQT was developed to assess divergent thinking, similar to the common and popular alternative uses task (AUT), which requires participants to generate alternative uses to objects (Acar & Runco, 2019). The UQT was conducted under timed conditions, and yielded two major scores, fluency, and originality. Fluency was scored as the total number of nonrepeated responses emitted (i.e., questions asked) while originality was scored by the use of norms and criteria reported by Torrance (1966). However, later versions of the TTCT removed the UQT from its battery of divergent thinking tests, which is perhaps part of the reason for the lack of research in the field of question asking and divergent thinking.

Creative people often question their assumptions and lead others to do the same in their search for answers (Sternberg & Williams, 1996). Albergaria-Almeida (2011) have argued for the importance of critical thinking skills in higher education and that creative abilities are directly related to questioning such that students ask questions that are coherent with their creativity levels. Albergaria-Almeida (2011) implemented several teaching and learning strategies in a chemistry and geology course, as a way of encouraging students' questioning. The authors found that lower creativity was associated with students that ask mainly closed, less complex questions which relate to simple facts and concepts, and higher creativity with those who can ask all kinds of questions including more complex, specialized higher-level questions which reveal working hypotheses and application of new knowledge (see also Acar et al., 2023).

Ronfard et al. (2018) examined current research on question asking in childhood by focusing on the epistemic function of questions—the use of questions to bridge a gap in knowledge or to resolve uncertainty. They highlight the great variability in the quality and quantity of the questions people ask, influenced by imprecise, poorly worded questions, as well as variablity in the number and precision of questions people ask. The authors argue that question asking is a powerful learning strategy, yet research on questions has been relatively sparse and isolated across several disciplines (but see Gottlieb, 2021; Nelson, 2005; Rothe et al., 2018; Sasson & Kenett, 2023). Therefore, an important question remains: What is the role of question asking in creativity? Based on Van Der Meij's (1994) componential analysis of student questioning, one possibility is that question asking is a critical mechanism in facilitating problem finding, the first stage of the creative process (Reiter-Palmon & Robinson, 2009).

Van Der Meij (1994) argues that when developing questions, the student must articulate and express the problem or perplexity they encounter. Van Der Meij (1994) claims it is quite likely that the student's questioning transition from perplexity to the formulation and expression of a question is a most difficult process and compares it to the subtle difference between problem finding and problem formulation. Whereas the first refers to finding the problem area, it is the second that enables the person to initiate a creative approach to the problem. Thus, question asking may be critical in facilitating problem finding behavior. However, it remains unclear how best to assess question asking ability. A possible approach to examine this issue is via the Bloom taxonomy.

The Bloom Taxonomy

Baloche (1994) argues that the teacher's skill in asking questions is very important to the development of their students' creative thinking skills and that higher level questions, which involve more complex and abstract ideas, such as creation of new topics and expression of opinions, are beneficial to fostering creativity. Thus, it seems that question level and complexity are an important aspect to consider in studying question asking. One possible approach to study question level is utilizing Bloom's taxonomy (Bloom et al., 1956). Bloom's taxonomy has been widely accepted as a guideline in designing cognitive learning objectives of differing levels (Adams, 2015; Goh et al., 2020; Omar et al., 2012). Specifically, the taxonomy includes six cognitive levels, which are hierarchically ordered from simple to complex. Previous studies analyzed Bloom taxonomy levels in a quantitative manner, where each question is assigned a Bloom's rating of one to six in accordance with each cognitive level (Oliver et al., 2004; Plack et al., 2007; Zheng et al., 2008). Krathwohl (2002) interprets the taxonomy as proceeding from simple to complex, in relation to increasing levels of the taxonomy. Thus, a quantitative, simple to complex rating of questions based on the Bloom taxonomy allows rating and classifying question complexity. Furthermore, it is assumed that the taxonomy represents a cumulative hierarchy; that is, mastery of each simpler category is a prerequisite to mastery of the next more complex one (Krathwohl, 2002). An updated version of Bloom's taxonomy includes the following levels, in ascending order of complexity (Krathwohl, 2002): remembering (retrieving relevant knowledge from long-term memory, for example when or how did X happen), understanding (determining the meaning of instructional messages, including oral, written, and graphic communication, e.g., how would you summarize…?), applying (carrying out or using a procedure in a given situation, e.g., how would you solve X using what you have learned), analyzing (breaking material into its constituent parts and detecting how the parts relate to one another, e.g., how can you make a distinction between…?), evaluating (making judgments based on criteria and standards, e.g., how would you prove or disprove…?), and creating (putting elements together to form a novel, coherent whole, e.g., what changes would you make to solve…?).

Bloom's taxonomy is widely used in education (Forehand, 2010; Granello, 2001); however, few attempts have been made to utilize it in measuring creativity in the context of cognitive psychology, and especially among adults (Grebin et al., 2020). Thus, we propose utilizing the Bloom taxonomy as a measure of question complexity to understand the role of question asking in general—and specifically more complex questions—in creativity. Furthermore, Smith (1970) and Roberts (1976) have found conflicting results on the relation between intelligence and cognitive abilities with Bloom taxonomy levels, with Smith (1970) finding support for the positive relation and Roberts (1976) finding none. As such, studying specific Bloom level of questions and their relation to creativity and intelligence may help elucidate the exact relation between cognitive abilities and question level. This can be achieved via relating question asking ability to individual differences in divergent thinking, the most commonly assessed cognitive component of creative thinking.

Divergent Thinking

One extensively studied aspect of creativity is divergent thinking. Divergent thinking tests are among the most commonly used in creativity research (Acar & Runco, 2019), conceived as a key cognitive foundation of creativity (Baer, 2012; Runco, 2010). Divergent thinking allows generating creative ideas by exploring many possible solutions or ideas. Some researchers suggest that this occurs in a spontaneous, but sometimes linear manner (Hass, 2017a, 2017b), such that many ideas are generated in an emergent fashion (McCrae, 1987; Runco, 1992). Because some of the resulting ideas are original, divergent thinking represents the potential for creative thinking (Runco, 1986). Additionally, previous meta-analyses found support that divergent thinking and creative achievement are modestly and positively related (Kim, 2008; Said-Metwaly et al., 2022).

One of the commonly used divergent thinking tasks is the AUT (Guilford, 1967). In the AUT, participants are presented with an object (e.g., a brick) and are required to think of as many uses as they can (Acar & Runco, 2019). Previous studies have shown that responses on such tasks typically follow a clear pattern in which participants initially generate many relatively common ideas, followed by increasingly more novel responses, generated at a slower pace (Bai et al., 2021; Beaty & Silvia, 2012). In the past two decades, divergent thinking has been strongly linked to intelligence (Batey et al., 2009; Lee & Therriault, 2013; Silvia, 2008, 2015), openness to experience (Batey & Furnham, 2006), and curiosity (Celik et al., 2016; Vidler & Karan, 1975). While question asking was originally included in the TTCT in assessing divergent thinking (Torrance, 1966), at some point it was removed from the TTCT, and research mostly focused on the AUT. Thus, it remains unclear whether the relations found between the AUT and intelligence, openness, and curiosity, are similar with question asking as assessed in the alternate questions task (AQT).

Additional Capacities Related to Creativity

There is ample reason to expect similar relations between question asking and additional cognitive factors (e.g., intelligence, openness, and curiosity), as found with the AUT. Studies have found an inconsistent relationship between question asking and intelligence (Roberts, 1976; Smith, 1970). Additionally, openness to experience, which has been linked to divergent thinking (Batey & Furnham, 2006), is strongly related to information-seeking behavior (Kenett et al., 2023; Woo et al., 2014), and as such might also play a role in question asking as related to creativity. Such is also the case for curiosity, as a motivating driver in creativity (Gross et al., 2020; Koutstaal et al., 2022), most likely realized by question asking (Kenett et al., 2023). Thus, several additional capacities that have been shown to be related to creativity may also play a critical role in question asking.

Fluid intelligence (*Gf*) is commonly defined as the ability to solve problems in unfamiliar domains using general reasoning methods (Avitia & Kaufman, 2014; Carroll, 1993; Cattell, 1963). Beaty and Silvia (2012) and Silvia (2015) have shown the strong relationship between *Gf* and creativity. Additionally, a latent *Gf* variable strongly predicted the creativity of response to unusual uses tasks (Silvia, 2008). Broad retrieval ability (*Gr*) has been found to have effects on divergent thinking that are distinct from the effects of *Gf*, such as ideational fluency and encoding ability (Silvia et al., 2013). *Gr* refers to the ability to retrieve information stored in long term memory, an ability that has been strongly linked to divergent thinking (Avitia & Kaufman, 2014). As such, we propose also testing *Gr* ability in this study as it may provide additional information on the relation between divergent thinking, creativity and question asking.

Openness to experience is one of the big-five personality traits; it is heritable, stable across adults, universal, and is related to divergent thinking and to creativity (Conner & Silvia, 2015; McCrae & Greenberg, 2014; Oleynick et al., 2017). In fact, in studies and in meta-analyses, openness to experience is perhaps the strongest positive predictor of creativity (Oleynick et al., 2017; Puryear et al., 2017).

A meta-analysis consolidated the results of studies of the association between curiosity and creativity and found a significant association between higher curiosity and greater creativity (Schutte & Malouff, 2020). But as Koutstaal et al. (2022) note, this meta-analysis only looked at self-reported creativity and curiosity, and not behaviorally assessed. Koutstaal et al. (2022) assessed curiosity by using a behavioral curiosity measure of novel question asking and the classic AUT measure of divergent thinking. The measure used was a novel Q&A curiosity task in which participants were presented with six brief factual statements about a variety of subjects, such as notable feats of mountain climbing. After each stimulus, participants were given the opportunity to ask—by typing into a text box—any questions that might naturally arise in relation to that stimulus. They found that originality of responses on the AUT significantly and positively correlated with the novelty of questions that participants asked on the curiosity Q&A task. We claim this provides further justification that question asking plays an important role in creativity, and mandates further exploration, perhaps even beyond the scope of curiosity.

The Present Research

The current study examines the role of question asking in creativity. In Study 1, we introduce the alternative questions task (AQT) to evaluate the relation between question asking ability and creative thinking. The AQT, which is based on the historic UQT by Torrance (1966), requires participants to generate creative and unusual questions about common objects such as a pillow or a pencil. The AQT differs from the classic UQT by focusing on question complexity—assessed via the Bloom taxonomy—and by capitalizing on recent advances in quantitative assessment of creativity (Kenett, 2019).

Subjective scoring of creativity, despite its strengths, has its limits; most notably, human raters do not always agree on what they find creative, and they are often asked to score thousands of responses—leading to rater fatigue and negatively impacting the reliability of their ratings (Forthmann et al., 2017; Kaufman, 2019; Silvia et al., 2008). Over the past several years, an increasing amount of research is using corpus-based computational linguistic models to compute the conceptual, or semantic, distance between participants' AUT responses and its object prompt (Beaty & Johnson, 2021; Dumas et al., 2021; Organisciak et al., 2023; Yu et al., 2023). Such a quantitative, objective, measure has been found to be reliably and positively related to subjective assessment of response novelty, negatively related to subjective assessment of response appropriateness (or usefulness), and positively related to subjective assessment of response creativity (Beaty & Johnson, 2021). Given that the AQT also assesses divergent thinking (Torrance, 1966), it is likely that such objective measures applied on the AUT can also be used in the AQT.

In the current study, we assess both AUT and AQT objective and subjective creativity scores. In addition, participant's AQT question responses were classified according to Bloom taxonomy levels by external raters. In addition, participants underwent *Gf*, *Gr*, openness, and curiosity assessments. In Study 2, we replicated and generalized the results of Study 1 by switching the stimuli between the AUT and AQT tasks to eliminate specific stimuli confounds. Thus, our research question is: What is the relation between question asking and creativity, and what is the role of complexity and additional capacities related to creativity in facilitating this relationship?

In both studies, we expected to find a relation between average Bloom question level and objective and subjective AQT and AUT creativity scores, that is, higher average Bloom level responses will be related to higher objective and subjective creativity scores (H1). We also theorized that higher levels in the taxonomy refer to more cognitively complex questions, while the lower levels represent lower complexity, with the average participant gravitating toward average Bloom level (H2). Additionally, we expected higher frequencies of lower Bloom level questions to correlate negatively with objective and subjective creativity scores, whilst higher frequencies of higher-level questions to be positively related (H3). Based on the UQT's origins as a measure of divergent thinking, we hypothesized a positive relation between the AQT and the AUT creativity scores (H4). Finally, we expected that performance in the AQT will also be related to *Gf*, *Gr*, openness, and curiosity scores as will performance on the AUT (H5).

The studies conducted here are the first of their kind and as such, exploratory in nature. Nevertheless, both studies were preregistered and available on as predicted (https://aspredicted.org/WBL_T1P). In addition to the preregistered analyses, we also ran an exploratory ad hoc analysis of response order and its effect on Bloom questions level scores across participants' responses on the AQT. We also analyzed the distribution of Bloom question levels across all items to examine the differences between each level.

Study 1

Study 1 aimed to directly examine the role of question asking in creative thinking. This was achieved via administrating the AQT. The AQT was inspired by the original UQT, and was constructed similarly to the AUT, where participants generate alternative uses for common objects. This was done under the assumption that this task structure would yield a valid and tested approach for generating fluent and novel question responses from participants (Torrance, 1966). In the AQT, participants are required to generate creative and unusual questions about common objects (pencil, pillow, and sock). Exploration of performance in this task will allow us to better understand the role of questioning abilities in creativity, as well as related cognitive capacities.

Method

Participants

Sample-size estimation for Study 1 was based on a priori power analyses using G*Power (Version 3.1.9.7; Faul et al., 2007) for linear multiple regression. This analysis determined that 118 participants would be necessary to achieve a *R* squared effect size of .30 with 99% power, α error probability of .05, and 10 tested predictors.

In total, 119 individuals recruited on Prolific Academic participated in the study for £4.6. Ten participants who did not provide answers for the AQT or AUT were excluded from the analysis. Thus, the final analyzed sample consisted of 109 participants (47.9% male, 50.4% female, 1.7% preferred not to say; $M_{age} =$ 26.1 years, $SD = 6.41$ years, mean years of formal education = 13.5 years, $SD = 5$ years). This research was approved by the Technion—Israel Institute of Technology Institutional Review Board.

Measures

The AUT

In the AUT (Acar & Runco, 2019; Guilford, 1967), participants are required to generate as many alternative uses as possible for a common object such as a pen, a brick, or a paperclip. For instance, alternative uses for a brick could include a bed riser, a place mat, or a weapon. AUT objects used in this study were taken from the suggested objects provided by Beaty et al. (2022) for the AUT (knife, purse, and clock). As in prior work (Runco et al., 2005; Said-Metwaly et al., 2020), participants were explicitly instructed to come up with as many original and creative uses for objects as they can. They were asked to come up with creative uses, which were defined in the study as uses that strike people as clever, unusual, interesting, uncommon, humorous, innovative, or different. Participants provided their responses on a single page with 30 available input fields and had access to their previous answers. Time limits were 2 min per object.

The AQT

The AQT, based on the UQT, requires participants to generate creative and unusual questions about three common objects in 2 min for each object. AQT objects were taken from the suggested items provided by Beaty et al. (2022) for the AUT (pencil, pillow, and sock). As in the AUT and in prior work (Said-Metwaly et al., 2020), participants were explicitly instructed to come up with as many original and creative questions for objects as they can. Creative questions were defined in the study as questions that strike people as clever, unusual, interesting, uncommon, humorous, innovative, or different. Participants provided their responses on a single page with 30 available input fields and had access to their previous answers. Time limits were 2 min per object. Although this seems like it might be a difficult task for participants, we decided not to give examples as to avoid biasing toward a specific level of question response, which would impact potential question level variance.

AUT and AQT Analyses

Fluency

Fluency was calculated as the average number of responses given across three items, computed separately for the AUT and AQT. Raw AUT and AQT responses were preprocessed by the first author to eliminate invalid or inappropriate responses, which consisted of incomplete sentences or irrelevant answers or symbols. A few of the responses to the AQT were largely incomplete, perhaps due to insufficient time while almost all AUT responses were appropriate. AQT responses that did not include a coherent, complete response in the form of a question, as instructed, were removed before analysis and scoring. Responses that contained adult language were also removed from the external scoring methods. Total excluded responses accounted for less than five percent of total responses.

Subjective Creativity Scoring

AQT responses were rated for creativity. The present research utilized snapshot scoring of divergent thinking tasks, in which the set of responses receives a single holistic rating (Runco & Mraz, 1992; Silvia et al., 2008). Online raters from Prolific Academic scored responses using a 1 (*not at all creative*) to 5 (*very creative*) scale (see Appendix A). Each object cue was rated by five different independent raters. Scores were averaged across the three AQT objects to create an average subjective creativity score for every participant. Raters were paid 4.6 pounds for their participation in the subjective creativity ratings. Raters who failed attention checks were excluded from the final data set. Reliability metrics for AQT subjective creativity scores were moderate (Koo & Li, 2016), and as follows: pencil ($N = 3$, $\alpha = .625$), pillow ($N = 3$, $\alpha = .56$), and sock ($N = 3$, $\alpha = .671$). These reliability scores correspond to previous studies (e.g., Beaty et al., 2023) using the same online crowdsourcing subjective rating of AUT responses approach (Hass et al., 2018).

Objective Creativity Scoring

Although creativity subjective scoring methods have proved useful, they have two inherent limitations—labor cost (raters typically code thousands of responses) and subjectivity (raters vary on their perceptions and preferences; Beaty & Johnson, 2021). As such, creativity researchers are increasingly using computational tools such as semantic distance to assess creativity (Kenett, 2018, 2019), as semantic distance provides a quantitative, objective, alternative to manual scoring by human raters (Beaty et al., 2022; Beaty & Johnson, 2021; Dumas et al., 2021; Organisciak et al., 2023). Semantic distance is based on computational linguistic models, that allow representing the similarity (or its inverse, distance) between concepts in multidimensional semantic spaces (Günther et al., 2019). In creativity research, such computational tools are utilized to quantitatively assess the semantic distance between the object and the open-ended response generated to it (such as in the AUT). Importantly, research has strongly demonstrated how such quantitative measures strongly relate to subjective ratings of idea originality, appropriateness (usefulness), and creativity (Beaty & Johnson, 2021; Dumas et al., 2021; Organisciak et al., 2023).

To conduct automated scoring of the AUT and AQT responses, we applied a recently developed automated creativity score, the maximum associative distance (MAD) scores (Yu et al., 2023). Within a response, MAD measures the semantic distance of the word that is maximally remote from the cue word to reflect response novelty and elaboration. Yu et al. (2023) found MAD to be more strongly correlated with human subjective creativity ratings than competing methods. In addition, MAD scores reliably predict external measures such as openness to experience. Even though MAD was developed specifically for the AUT, its high correlation with human creativity ratings and emphasis on measuring divergent thinking, combined with the AQT's origins as a historically divergent thinking task, make MAD a promising potential tool for measuring creativity in the AQT.

To evaluate the objective creativity of a response, we computed MAD scores for the AUT and AQT stimuli (e.g., pencil) and participants' responses by computing the distance between the vectorized word representations in a high dimensional vector space (semantic space) using the MAD python package as done by Yu et al. (2023). MAD scores were averaged across the three items to obtain average AQT MAD scores and average AUT MAD scores for each participant as their objective creativity scores.

Bloom Hierarchy

AQT responses were scored for their respective Bloom level (from one to six: remember, understanding, applying, analyzing, evaluating and creating). The revised edition of Bloom's taxonomy (Bloom et al., 1956; Krathwohl, 2002) was used. Online raters from Prolific Academic were instructed on the Bloom taxonomy levels and rated AQT responses by assigning the Bloom level they ascertained from the relevant AQT response (see Appendix B). Rating instructions included an explanation of the types of questions asked for each Bloom level, alongside key terms related to each Bloom level. Each object cue was rated by 10 different independent raters. Raters who failed attention checks or gave incomplete ratings were excluded from the final data set. Reliability metrics for AQT objects on their Bloom level ratings were overall good (Koo & Li, 2016) as follows: pencil ($N = 4$, $\alpha = .752$), pillow ($N = 3$, $\alpha = .727$), and sock ($N = 10$, $\alpha = .768$). Average Bloom scores were then calculated as the average level score of all participant's responses across all AQT objects. Examples of responses corresponding to each bloom level as rated by external trained raters for the cue pillow are presented in Table 1.

Intelligence

Participants completed intelligence tasks, assessing lower-order facets of general intelligence: fluid intelligence (*Gf*) and broad retrieval ability (*Gr*). *Gf* tasks include: (a) a number series task (18 items, 6 min), which presents sequences of numbers that change based on a rule and asks participants to select the next sequence (Thurstone, 1938) and (b) a series completion task from the Culture Fair Intelligence Test (16 items, 4 min), which presents sequences of three changing images (small line drawings) and asks participants to select the next image that fits the rule governing their change (Cattell & Cattell, 1961/2008). Scores of both *Gf* tasks were averaged into one average *Gf* score. *Gr* tasks include a category fluency task: listing animals (2 min; Ardila et al., 2006).

Personality

Personality was measured using the big-five factor model, including facets, based on the NEO-PI-3 questionnaire (Costa & McCrae, 1992). We measured only openness to experience/intellect (23 items, 7 min) as it has been found to correlate with creativity (Oleynick et al., 2017). Internal consistency was acceptable ($\alpha = .770$).

Curiosity

All participants completed a recently developed short 22-items questionnaire that assesses various aspects of curiosity. These 22 items were identified and selected from the following existing questionnaires: the five-dimensional curiosity–social curiosity, workplace curiosity–organizational curiosity, I/D-type curiosity–intellectual cognition, information seeking, and perceptual curiosity (Collins et al., 2004; Kashdan et al., 2020; Litman, 2008; Litman & Jimerson, 2004; Litman & Spielberger, 2003). Reliability for the adapted scale was acceptable ($\alpha = .702$).

Table 1
Examples of Question Responses Corresponding to Each Bloom Level as Rated by External Trained Raters for the Object Pillow

Bloom level	Example
1. Remembering—these types of questions address terms, facts, and details without necessarily understanding the concept.	Why isn't it hard?
2. Understanding—these questions test the ability to summarize and describe in our own words without necessarily relating it to anything.	What is it for?
3. Applying—these questions involve applying or transferring learning to our own life or to a context different than one in which we learned.	What's the most convenient pillow?
4. Analyzing—these questions are about breaking material into parts, describing patterns and relationships among parts, to subdivide information and to show how it is put together.	Do you like to put your hand inside the pillowcase while sleeping?
5. Evaluating—evaluation questions involve developing opinions and making value decisions about issues based on specific criteria.	Do you prefer to have the buttons of the pillowcase to be on the left or right?
6. Creating—these questions involve creating something new by using a combination of ideas from different sources to form a new whole.	Would it be possible to create biodegradable pillows, since after years of use, they're gross and unrecyclable?

Statistical Analysis

We conducted Pearson's correlation analyses to examine the relation between Bloom levels, and subjective creativity and objective creativity scores. Average Bloom scores were calculated as the average level score of all the participant's responses across the AQT object prompts. The average Bloom score is a measure that we hypothesized to be positively correlated to subjective and objective creativity such that higher average participant Bloom scores correlate with higher objective and subjective creativity scores. We also tested the correlation between the frequency of each individual Bloom level and the other variables in our data set—intelligence, personality, and curiosity, and our control variables of sex, age, and level of education. This individual Bloom level analysis is meant to further differentiate between high (more complex) and low (less complex) Bloom level questions. We hypothesized here that lower Bloom levels will correlate negatively with creativity scores, medium levels will not be correlated, and high levels will be positively correlated. Finally, we computed multiple linear regression models that tried to predict AUT and AQT objective creativity scores, based on the other remaining variables that were tested and known in the literature to correlate with divergent thinking tasks. In addition to the preregistered analyses described above, we also ran an analysis of variance (ANOVA) of response order and its effect on Bloom question level scores across participants' responses on the AQT. We also analyzed the distribution of Bloom question levels across all items to examine the differences between each level using a repeated measures ANOVA.

Procedure

Participants were asked to endorse a consent form to indicate they agree to participate in the study. Upon consent, participants received an explanation of the course of the experiment. Participants then generated questions in the AQT for three randomly ordered objects (pencil, pillow, and sock) and then answered in the following order—a *Gf* intelligence task, an openness to experience personality questionnaire, a curiosity questionnaire, and lastly a *Gr* semantic fluency task. Next, they were asked to generate alternate uses in the AUT for three different objects (knife, purse, clock). Finally, they filled out a demographic questionnaire. We excluded only participants who failed to finish the experiment or who gave incomplete answers or skipped more than a few questions or failed to provide responses for the AQT or AUT. Attention checks were randomly placed during the questionnaire parts of the experiment as questions instructing participants to select a certain answer.

Results

Comparing AQT and AUT Performance

We began our analysis by preprocessing AUT and AQT raw responses for subsequent planned analyses, by removing invalid or inappropriate responses. A few responses to the AQT were largely incomplete, perhaps due to insufficient time while almost all AUT responses were appropriate. Responses that contained adult language were also removed. Total excluded responses accounted for less than five percent of total responses. To assess the difficulty of the AQT compared to the AUT task, we began by examining participants fluency performance in the AQT compared to the AUT, alongside their additional behavioral assessment (Table 2). This was done in order to dispel concerns that perhaps the AQT responses were limited due to time constraints. This analysis revealed no significant differences in fluency between the AUT ($M = 5.192$, $SD = 2.071$) and AQT ($M = 5.311$, $SD = 2.297$), $t(108) = 0.539$, $p = .591$, $d = -0.05$. Thus, the time limits were sufficient for both the AQT and AUT. Furthermore, at the end of the study we gave participants the opportunity to provide feedback on the difficulty of the study. Qualitative analysis of these responses revealed that almost all participants reviewed the task as enjoyable and interesting but not too difficult. The results also showed that participants differed in their objective creativity scores, generating more semantically distant responses to the AQT ($M = 0.904$, $SD = 0.01$) compared to the AUT ($M = 0.808$, $SD = 0.05$), $t(108) = 18.95$, $p < .001$, $d = 1.8$. We then proceeded to analyze all tasks conducted in this study. Table 2 provides descriptive information for all variables examined in this study.

Correlations of Main Study Variables

Next, we examined the relationship of the AQT with all other assessed variables using Pearson's correlation analysis (Figure 1;

Table 2
Descriptive Information for Study Variables

Variable	*M*	*SD*	Min	Max
AUT fluency	5.192	2.071	1.67	12.67
AUT objective creativity	0.808	0.054	0.64	0.90
AQT fluency	5.311	2.297	1	14
AQT objective creativity	0.904	0.015	0.86	0.94
AQT subjective creativity	2.277	0.346	1.54	3.31
AQT bloom score	2.793	0.458	1.74	3.62
Gf	6.825	2.00	2.5	10.5
Curiosity	3.541	0.456	2.18	4.77
Openness to experience	70.733	8.261	47	92
Semantic fluency *Gr*	25.76	8.56	6	50
Age	26.222	6.433	18	51
Years of formal education	13.532	5.154	10	23

Note. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability.

correlation table can be found in Appendix C). Additionally, to further understand the effect of question complexity on creativity, we separately examined Bloom levels one to six's relation to objective and subjective AQT creativity (Figure 1). Bloom level-based scores are the frequency of answers given by a participant for each one of the six Bloom taxonomy levels.

Average Bloom levels (the average Bloom question level of all participants' AQT responses) were significantly positively correlated with AQT objective creativity scores, $r(109) = .510$, $p < .001$, and AQT subjective creativity scores, $r(109) = .552$, $p < .001$. Bloom average questions levels were negatively correlated with AQT fluency, $r(109) = -.249$, $p = .009$. AQT objective creativity scores were also significantly positively correlated with AUT objective creativity scores, $r(109) = .247$, $p = .01$, AQT subjective creativity scores, $r(109) = .543$, $p < .001$, and negatively correlated with AQT fluency, $r(109) = -.333$, $p < .001$. Furthermore, AQT subjective creativity scores were significantly positively correlated with *Gf*, $r(109) = .238$, $p = .013$, and negatively correlated with AQT fluency, $r(109) = -.343$, $p < .001$, while AUT objective creativity scores were positively correlated with *Gf*, $r(109) = .249$, $p = .009$.

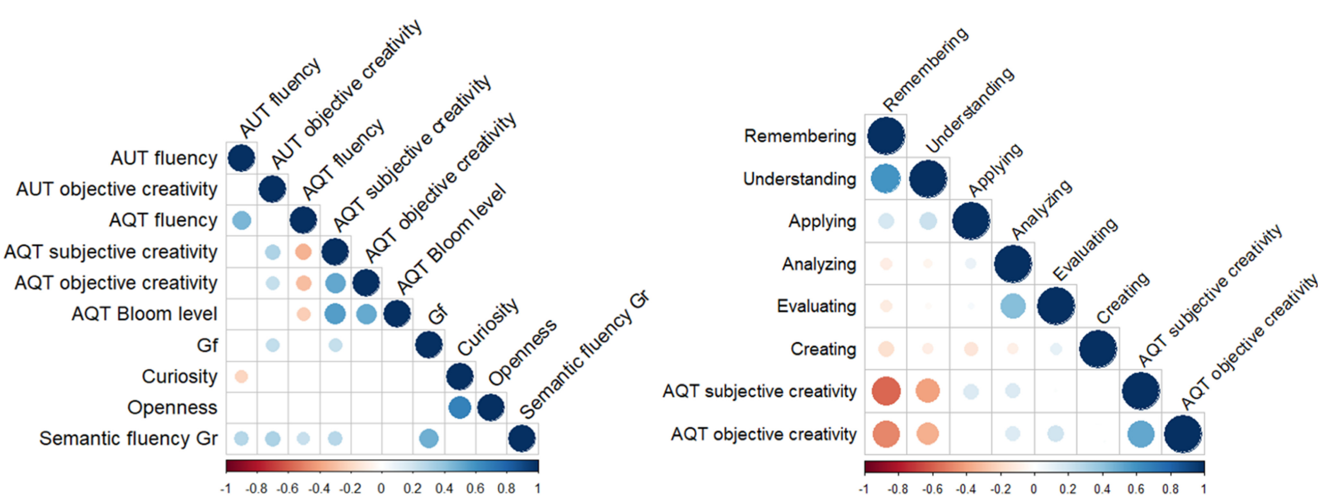
Although not in the preregistration, as a follow up test we analyzed the distribution of Bloom question levels across all items in order to examine the differences between Bloom levels and to test the effect of fluency on each level. The average distribution of question levels across all items tested were Bloom Level 1: 14.66%, Bloom Level 2: 24.45%, Bloom Level 3: 35.87%, Bloom Level 4: 17.30%, Bloom Level 5: 7.33%, and Bloom Level 6: 0.36%. Additionally, a Kolmogorov–Smirnov test indicated that average Bloom level scores follows a normal distribution, $D(108) = 0.042$, $p = .200$.

We then examined specific Bloom level response frequency for each participant. A Greenhouse–Geisser repeated measures within-subject ANOVA revealed that frequencies of responses differed significantly for individual Bloom levels, $F(5, 540) = 90.928$, $p < .001$, $\eta_p^2 = .457$. Post hoc Bonferroni corrected paired-samples *t*-test analyses revealed that all comparisons were significantly different (all *ps* $< .05$), except the comparison between Bloom Level 1 ($M = 2.238$, $SD = 2.55$) and Bloom Level 4 ($M = 2.64$, $SD = 2.01$), $t(109) = 0.404$, $p = 1$.

Individual Bloom Level Analysis

When analyzing the data according to individual Bloom levels, calculated as the frequency of questions for each stage (Figure 1), Bloom Level 1, remembering, was significantly negatively correlated with AQT objective creativity scores, $r(109) = -.417$, $p < .001$, as well as with AQT subjective creativity scores, $r(109) = -.492$, $p < .001$. Bloom Level 2, understanding, was also significantly negatively correlated with AQT objective creativity scores, $r(109) = -.196$,

Figure 1
Pearson's Correlation Analysis of AQT



Note. Left: General analysis of AQT with all other variables; Right: Analysis of individual Bloom levels and AQT objective and subjective creativity scores. The size of the circles represents the strength of the relation while the colors represent the direction of the relation (darker blue/dark gray in print for positive and darker red/light gray in print for negative). Only significant relations are shown. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability. See the online article for the color version of this figure.

$p = .043$, and negatively correlated with AQT subjective scores, $r(109) = -.248$, $p = .01$. Bloom Level 3, applying, was significantly positively correlated with AQT subjective creativity scores, $r(109) = .281$, $p = .003$, and nonsignificant with AQT objective creativity scores, $r(109) = .079$, $p = .417$. Bloom Level 4,, analyzing, was significantly and positively correlated with AQT objective creativity scores, $r(109) = .293$, $p = .002$, and significantly and positively correlated with AQT subjective creativity scores, $r(109) = .318$, $p < .001$. Bloom Level 5, evaluating, was significantly and positively correlated with AQT objective creativity scores, $r(109) = .304$, $p < .001$. Bloom Level 6, creating was marginally significantly correlated with AQT objective creativity scores, $r(109) = .188$, $p = .052$. This might be because it was less than 1% of AQT responses ($N = 6$) and thus suffers from low statistical power due to small sample size.

Additionally, we also analyzed the relation between individual Bloom levels, calculated as the frequency of questions for each level, and Fluid intelligence G*f*, curiosity, openness, semantic fluency *Gr* and AUT objective creativity scores. Bloom Levels 1, 2, 4, 5, and 6 were nonsignificant across all variables, while Bloom Level 3 was positively correlated with semantic fluency *Gr*, $r(109) = .209$, $p = .029$.

Predicting AQT and AUT Creativity Scores

Furthermore, in order to try to predict AQT and AUT creative performance and to examine whether they predict each other, a multiple linear regressions was conducted with AQT and AUT objective creativity scores as the dependent variable (Table 3, Figure 2). The variables entered were average Bloom level scores, openness and curiosity scores, *Gf* scores, and semantic fluency *Gr* scores. Since subjective and objective creativity scores were found to be highly correlated in the literature (Beaty & Johnson, 2021; Yu et al., 2023), AQT and AUT subjective creativity scores were not entered in the model to avoid problems of multicollinearity. None of the control variables (age, sex, and years of formal education) were significantly correlated with our dependent variables. For AQT objective creativity scores, a multiple regression analysis was applied to predict AQT originality scores from openness, *Gf*, semantic fluency *Gr*, AUT objective creativity scores, curiosity, and Bloom average scores along with age, sex, and years of formal education as control variables. These variables significantly predicted AQT objective creativity scores, $F(9,\ 95) = 4.851$, $p < .001$, $R^2 = .315$. Only AUT objective creativity scores ($p = .044$), and average Bloom scores ($p < .001$) were statistically significant predictors.

For AUT objective creativity scores, a multiple regression analysis was applied to predict AUT creativity scores from openness, *Gf*, semantic fluency *Gr*, AQT objective creativity scores, curiosity, and Bloom average scores along with age, sex, and years of formal education as control variables. These variables significantly predicted AUT creativity scores, $F(9,\ 95) = 2.424$, $p = .016$, $R^2 = .187$. Only semantic fluency *Gr* ($p = .009$), and AQT objective creativity scores ($p = .044$) were statistically significant predictors. There were no issues of multicollinearity in either analysis. Variance inflation factors (VIFs) were used to help detect multicollinearity. VIF is used to detect the severity of multicollinearity in regression analysis by measuring the number of inflated variances caused by multicollinearity. When VIF is higher than 10, there is significant multicollinearity that needs to be corrected (Curto & Pinto, 2011), whereas in our data all VIF values were much smaller.

Serial Order Effect

Although it was not included in the study preregistration but based on the serial order effect found in AUT (Beaty & Silvia, 2012), we decided to test exploratory ad hoc effects of response order on average Bloom questions level scores across participants' responses on the AQT (Figure 3). This was accomplished by testing for significant differences in average Bloom scores across participants' ordered responses, that is, testing if subsequent responses after their first response differed in Bloom score, etc. This was done to better understand possible response order effects, which may clarify creative question formulation and possible directions for further study into manipulating question complexity. We conducted this exploratory analysis on participants' first five AQT responses, based on the mean fluency of AQT responses which was close to five responses, and because sample size was much lower after the fifth response. A one-way ANOVA revealed no significant difference between response

Table 3

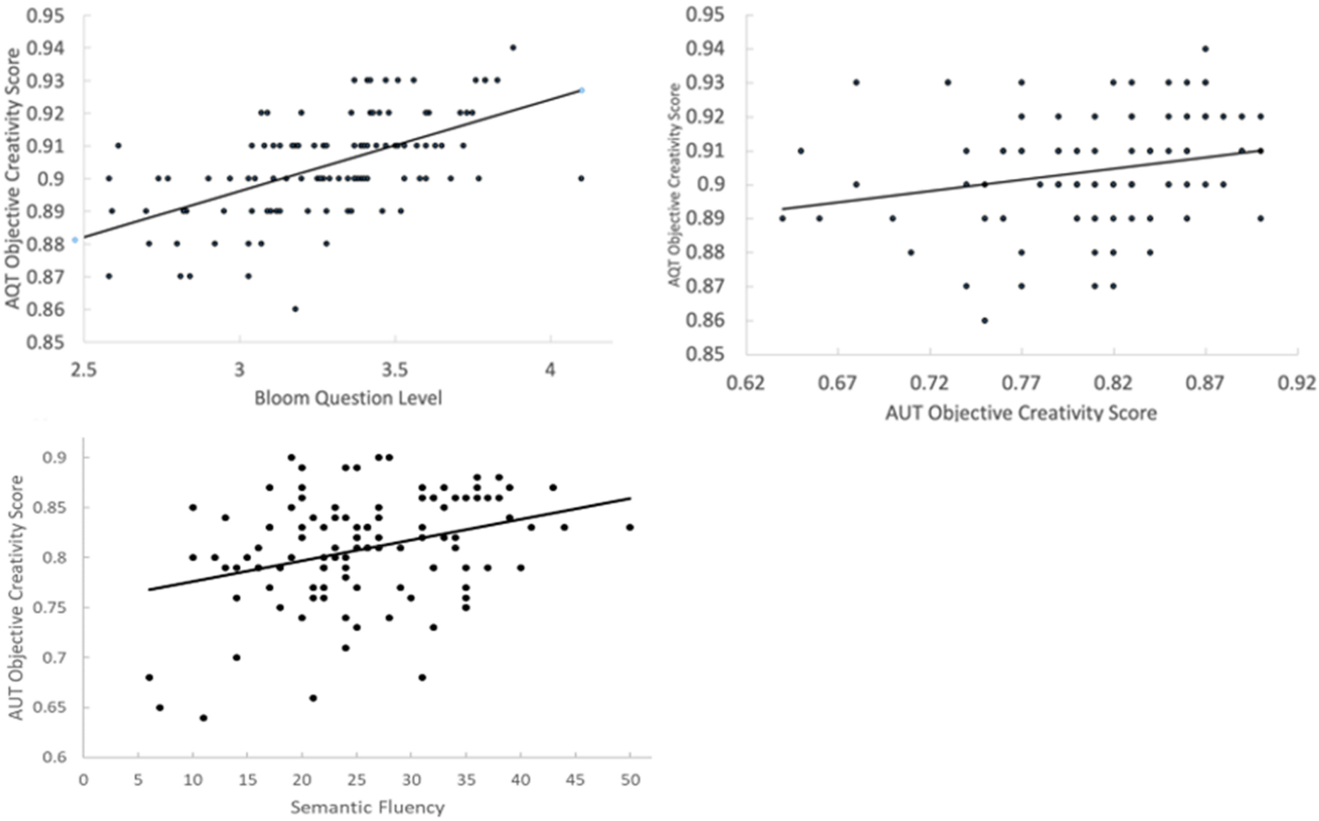
Summary of Multiple Regression Analysis for Variables Predicting AQT and AUT Objective Creativity (Nonsig. Control Variables Are Not Shown)

Effect	Estimate (*B*)	*SE*	95% CI (LL)	95% CI (UL)	Significance	VIF
AQT objective creativity						
Bloom taxonomy average	.016	.003	0.11	0.022	<.001	1.051
AUT objective creativity	.05	.025	0.001	0.099	.044	1.2
Openness	.000	.000	−0.001	0.000	.445	1.937
Gf	.000	.001	−0.001	0.002	.550	1.606
Semantic fluency *Gr*	.002	.001	0.000	0.003	.794	1.573
Curiosity	.004	.172	−0.002	0.013	.137	2.079
AUT objective creativity						
AQT objective creativity	.842	.412	0.025	1.659	.044	1.594
Gf	.004	.003	−0.002	0.011	.178	1.611
Semantic fluency *Gr*	.002	.001	0.000	0.003	.009	1.358
Curiosity	.016	.066	−0.024	0.040	.471	2.123
Bloom taxonomy average	.009	.013	−0.016	0.34	.461	1.389

Note. CI = confidence interval; LL = lower limit; UL = upper limit; VIF = variance inflation factor; AQT = alternative questions task; AUT = alternative uses task; *Gf* = fluid intelligence; *Gr* = retrieval ability.

Figure 2

Scatter Plots of the Relation of AQT Objective Creativity Scores With (in Clockwise Order From Top Left) Bloom Question Level, AUT Objective Creativity Scores, and AUT Objective Creativity Scores and Semantic Fluency Gr



Note. Circles represent average participants scores, trendline represents correlations described above. AQT = alternative questions task; AUT = alternative uses task. See the online article for the color version of this figure.

Bloom level, $F(4, 1383) = 0.484$, $p = .748$, $\eta_p^2 = .001$, meaning a serial effect for response order was not found.

Discussion

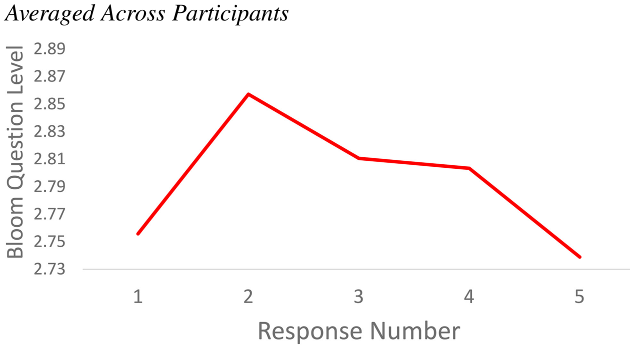
Study 1 aimed to directly examine the role of question complexity on creative thinking. This was achieved via the AQT—a revised version of Torrance's (1966) UQT—and testing the questions generated by participants on subjective creativity and objective measures of creativity, in addition to rating their respective levels in the Bloom taxonomy.

The results showed that participants differed in their objective creativity scores, generating more semantically distant responses to the AQT than in the AUT, while fluency was not significantly different. In addition, we found significant positive relations between AQT objective and subjective creativity scores, and both were also positively related with average Bloom level score. Secondly, we found a positive significant relationship between AQT and AUT objective creativity scores, which support the notion that these are related but separate measures. For predicting AUT objective creativity scores, a multiple regression analysis found that semantic fluency *Gr* and AQT objective creativity scores were significant predictors. For predicting AQT objective creativity scores, a multiple regression analysis found that AUT objective creativity scores and average Bloom scores were significant predictors.

Interestingly, individual Bloom questions levels were differently related to their objective creativity scores: Whereas lower levels such as remembering and understanding were negatively related to AQT objective creativity scores, higher levels such as analyzing and evaluating were positively related to AQT objective creativity

Figure 3

Line Chart of Average Bloom Levels by Order of Response—Averaged Across Participants



Note. See the online article for the color version of this figure.

scores. Regarding the relation between individual Bloom levels, and our cognitive measures of intelligence, curiosity, openness, semantic fluency *Gr* and AUT objective creativity scores—Bloom Level 3 was positively correlated with semantic fluency *Gr*. All other levels were not significantly related. An ad hoc ANOVA analysis of response order effects on Bloom level showed that a serial effect for response order was not found.

Overall, the results of Study 1 highlight the potential of the AQT coupled with Bloom's taxonomy in studying creativity and question asking. However, since Study 1 was exploratory in nature, we conducted Study 2 to replicate and generalize these results.

Study 2

Study 2 aimed to generalize and replicate Study 1, by switching the AUT and AQT objects used in Study 1. This was done to avoid possible confounds caused by item specific effects in the AQT, and to extend the validity and reliability of our findings.

Materials and Method

Participants

In line with Study 1 calculated sample size and based on resource constraints, 120 individuals were recruited on Prolific Academic, and participated in the study for £4.6. Six participants who didn't give answers for the AQT or AUT were excluded from the study bringing the final sample to 114 (49.1% male, 47.5% female, 3.4% prefer not to say; $M_{\text{age}} = 25.5$ years [$SD = 7.24$ years], mean years of formal education = 14.4 years [$SD = 4.4$]). This study was approved by the Technion—Israel Institute of Technology Institutional Review Board.

Method

Methods were the same as in Study 1. However, AQT objects were switched with AUT objects so that now the AQT objects were knife, purse, and clock, and the AUT objects were pillow, sock, and pencil. Five online raters from Prolific Academic scored responses using a 1 (*not at all creative*) to 5 (*very creative*) scale. Each object prompt was rated by different independent raters. Scores were averaged across the three AQT objects to create an average subjective creativity score for every participant. Raters who failed attention checks or gave incomplete ratings were excluded from the final data set. Reliability metrics for AQT subjective creativity scores were moderate (Koo & Li, 2016) as follows: knife ($N = 5$, $\alpha = .702$), purse ($N = 3$, $\alpha = .63$), and clock ($N = 4$, $\alpha = .61$).

Similar to Study 1, online raters from Prolific Academic were instructed on the Bloom taxonomy levels and rated AQT responses by assigning the Bloom level they ascertained from the relevant AQT response. Each object prompt was rated by 10 different independent raters. Raters who failed attention checks or gave incomplete ratings were excluded from the final data set. AQT Bloom reliability for Study 2 was good (Koo & Li, 2016) as follows: knife ($N = 6$, $\alpha = .732$), purse ($N = 3$, $\alpha = .706$), and clock ($N = 9$, $\alpha = .737$).

Statistical Analysis

Statistical analysis was the same as in Study 1.

Procedure

The procedure was the same as in Study 1.

Results

Comparing AQT and AUT

Firstly, like in Study 1, we calculated fluency scores by analyzing raw AUT and AQT responses and eliminating invalid or inappropriate responses. AQT responses that did not include a coherent, complete response in the form of a question, as instructed, were removed before analysis and scoring. Total excluded responses accounted for less than five percent of total responses. Then, we examined participants' performance in the AQT compared to the AUT, alongside their additional behavioral assessment (Table 4). Similar to Study 1, we found that differences in fluency between AUT ($M = 5.967$, $SD = 2.328$) and AQT ($M = 5.485$, $SD = 2.385$) were not significant, $t(112) = 1.401$, $p = .164$, $d = 0.128$. Similar to Study 1, we find significant differences between the objective creativity scores, with more semantically distant responses generated to the AQT ($M = 0.89$, $SD = 0.01$) compared to the AUT ($M = 0.82$, $SD = 0.04$), $t(112) = 17.544$, $p < .001$, $d = 1.6$. Table 4 provides descriptive information for all variables analyzed in Study 2.

Correlations of Main Study Variables

We then examined the relation of the AQT with all other assessed variables by a Pearson's correlation analysis (Figure 4; correlation table can be found in Appendix D). Additionally, to further understand the effect of question complexity on creativity, we separately examined Bloom Levels 1 to 6's relation to AQT objective and subjective creativity scores (Figure 4). Bloom level-based scores are the frequency of answers given by a participant for each one of the six Bloom taxonomy levels.

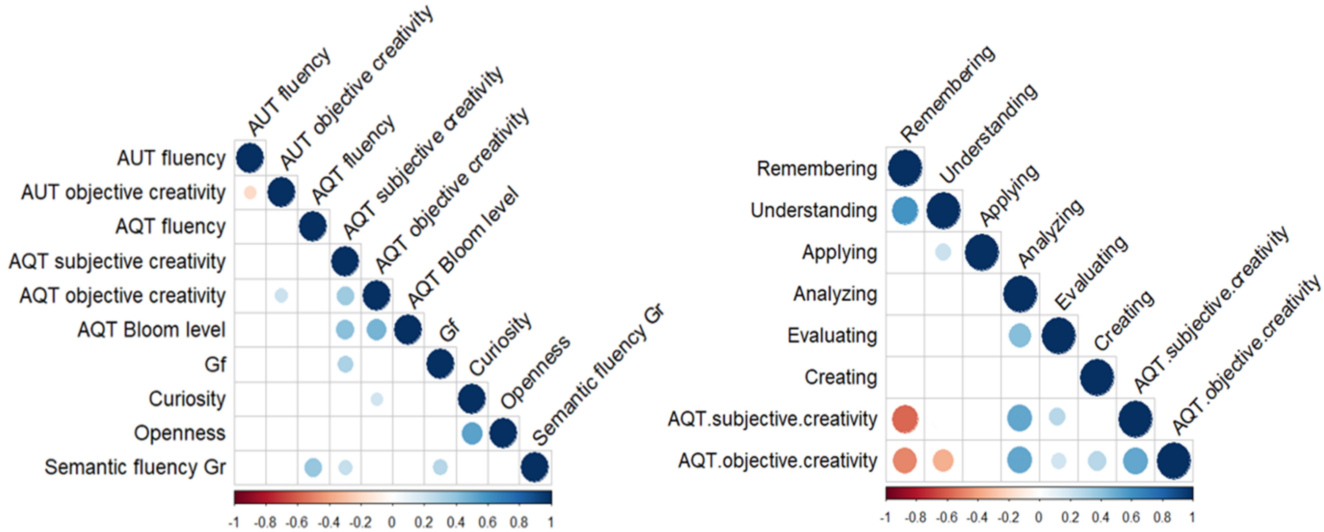
Average Bloom levels were significantly positively correlated with AQT objective creativity, $r(114) = .470$, $p < .001$, and positively correlated with AQT subjective creativity, $r(114) = .413$, $p < .001$. AQT objective creativity was positively correlated with AUT objective creativity, $r(113) = .198$, $p = .035$. Furthermore, AQT subjective creativity was positively correlated with *Gf*, $r(114) = .311$, $p < .001$ and *Gr*, $r(114) = .239$, $p = .011$. Contrary

Table 4
Descriptive Information for Study Variables

Variable	*M*	*SD*	Min	Max
AUT fluency	5.967	2.328	1.67	12.67
AUT objective creativity	0.821	0.043	0.64	0.90
AQT fluency	5.485	2.385	1	14
AQT objective creativity	0.891	0.012	0.86	0.94
AQT subjective creativity	2.522	0.338	1.54	3.31
AQT Bloom score	2.735	0.379	1.9	3.75
Gf	7.294	1.875	3.5	11.5
Curiosity	3.548	0.413	2.18	4.77
Openness to experience	71.09	7.890	47	92
Semantic fluency *Gr*	27.25	8.487	3	50
Age	25.5526	7.24183	18	50
Years of formal education	14.4386	4.40642	12	21

Note. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability.

Figure 4
Pearson's Correlation Analysis of All Variables



Note. Left: General analysis of AQT with all other variables; Right: Analysis of individual Bloom levels and AQT objective and subjective creativity scores. The size of the circles represents the strength of the relation while the colors represent the direction of the relation (darker blue/dark gray in print for positive and darker red/light gray in print for negative). Only significant relations are shown. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability. See the online article for the color version of this figure.

to Study 1, AQT subjective creativity was not significantly correlated with AQT fluency.

Although it was not in the preregistration, we analyzed the distribution of Bloom question levels across all items to examine the differences between each. The average distribution of question levels across all items tested were similar to Study 1. A Kolmogorov–Smirnov test indicated that average Bloom level scores again follow a normal distribution, $D(112) = 0.051$, $p = .2$.

We then examined specific Bloom level response frequency for each participant. A Greenhouse–Geisser repeated measures ANOVA showed that frequencies of responses differed significantly for individual Bloom levels, $F(5, 555) = 128.851$, $p < .001$, $\eta_p^2 = .537$. Similarly, to Study 1, post hoc Bonferroni corrected paired samples *t*-test analyses revealed that all comparisons were significantly different (all $ps < .01$), except for one contrast: Bloom Level 1 ($M = 2.491$, $SD = 2.28$) compared to Bloom Level 4 ($M = 2.571$, $SD = 1.925$), $t(112) = 0.08$, $p = 1$, $d = 0.024$.

Individual Bloom Level Analysis

When analyzing the data according to the individual Bloom levels (Figure 4), Similarly to Study 1, Bloom Level 1, remembering, was significantly negatively correlated with AQT objective creativity scores, $r(114) = -.335$, $p < .001$, and negatively correlated with AQT subjective creativity scores, $r(114) = -.238$, $p < .001$. Bloom Level 2, understanding, was significantly negatively correlated with AQT objective creativity scores, $r(114) = -.260$, $p = .006$, and positively correlated with AQT fluency, $r(114) = .350$, $p < .001$. Bloom Level 3, applying, was positively correlated with AQT fluency, $r(114) = .379$, $p < .001$. Bloom Level 4, analyzing, was positively correlated with AQT objective creativity scores, $r(114) = .429$, $p < .001$, and AQT subjective creativity, $r(114) = .408$, $p < .001$. Bloom Level 5, evaluating, was significantly positively correlated with AQT objective creativity scores, $r(114) = .191$, $p = .044$, AQT subjective creativity scores, $r(114) = .203$, $p = .032$. Bloom Level 6, creating, was positively correlated with AQT objective creativity scores, $r(114) = .260$, $p = .006$.

Additionally, we analyzed the relation between individual Bloom levels, calculated as the frequency of questions for each stage, and our cognitive measures of intelligence, curiosity, openness, semantic fluency *Gr* and AUT objective creativity scores. Bloom Level 1 was negatively correlated with curiosity, $r(113) = -.276$, $p = .003$, and openness to experience $r(105) = -.214$, $p = .023$; Bloom Level 2 was negatively correlated with curiosity, $r(113) = -.241$, $p = .011$; Bloom Level 3 was nonsignificant across all variables; Bloom Level 4 was marginally positively related to AUT objective creativity, $r(109) = .184$, $p = .052$, and significantly positively related to semantic fluency *Gr* $r(109) = .258$, $p < .001$ and *Gf* $r(109) = .187$, $p = .05$, while interestingly Bloom Level 5 was also modestly negatively related to curiosity, $r(112) = -.210$, $p = .026$. Bloom Level 6 was nonsignificant across all variables.

Predicting AQT and AUT Creativity Scores

A multiple linear regression was conducted with AQT and AUT objective creativity scores as the dependent variable (Table 5, Figure 5). The control variables—age, sex, and years of formal education were not significantly correlated with the dependent variables. For AQT objective creativity scores, a multiple regression analysis was applied to predict AQT creativity scores from openness, *Gf*, semantic fluency *Gr*, AUT objective creativity scores, curiosity, and Bloom average scores along with age, sex and years of formal education as control variables. These variables significantly predicted AQT objective creativity scores, $F(9, 95) = 5.543$, $p < .001$,

Table 5

Summary of Multiple Regression Analysis for Variables Predicting AQT and AUT Objective Creativity (Nonsig. Control Variables Are Not Shown)

Effect	Estimate (*B*)	*SE*	95% CI (LL)	95% CI (UL)	*p*-value	VIF
AQT objective creativity						
AUT objective creativity	.052	.025	0.001	0.102	.045	1.129
Openness	.000	.000	0.000	0.000	.905	1.535
Gf	.000	.001	−0.001	0.001	.942	1.163
Semantic fluency *Gr*	.000	.000	0.000	0.000	.384	1.130
Curiosity	.003	.003	−0.002	0.009	.256	1.109
Bloom taxonomy average	.014	.003	0.008	0.019	>.001	1.068
AUT objective creativity						
AQT objective creativity	.774	.338	0.104	1.4	.024	1.087
Openness	.000	.000	−0.001	0.001	.728	1.516
Gf	−.006	.003	−0.013	0.000	.042	1.320
Semantic fluency *Gr*	.001	.001	0.000	0.002	.122	1.169
Curiosity	.011	.012	−0.013	0.034	.373	1.49
Bloom taxonomy average	.012	.013	−0.013	0.038	.345	1.2

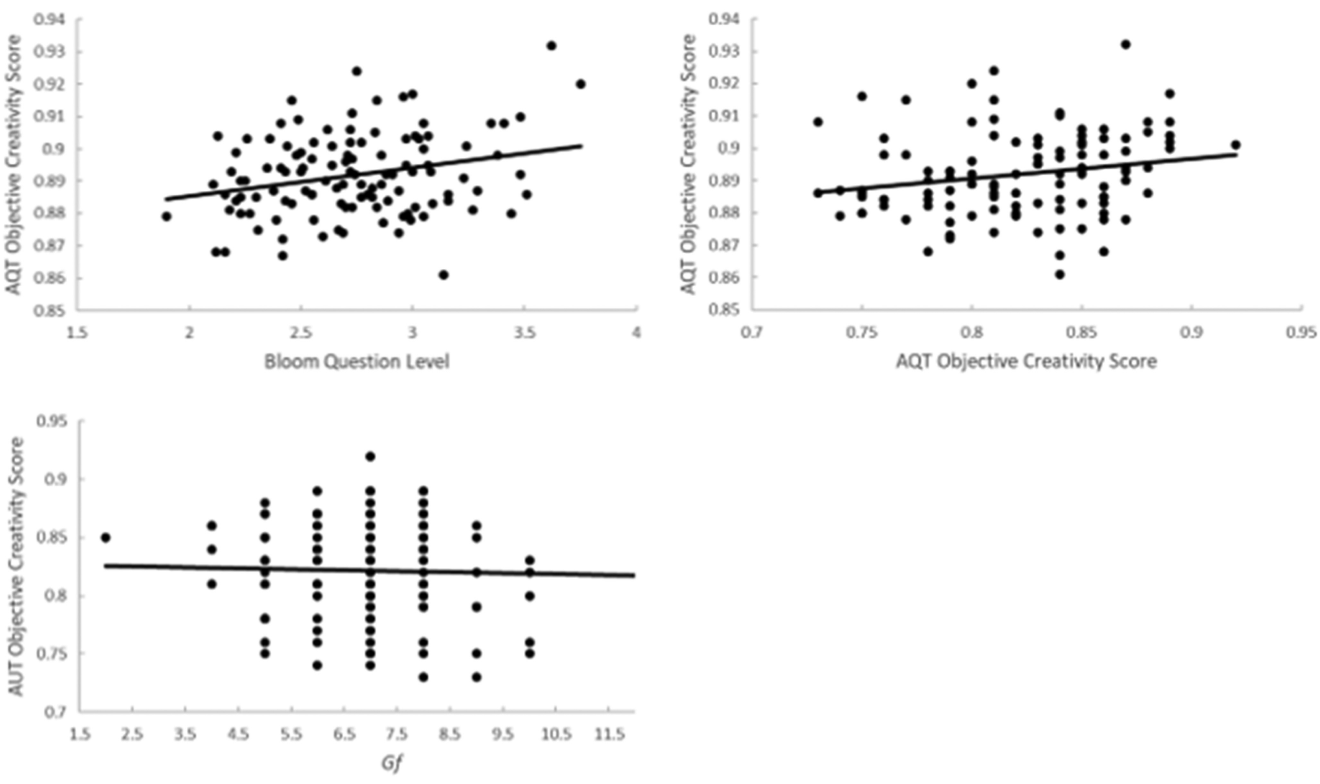
Note. CI = confidence interval; LL = lower limit; UL = upper limit; VIF = variance inflation factor; AQT = alternative questions task; AUT = alternative uses task; *Gf* = fluid intelligence; *Gr* = retrieval ability.

$R^2 = .286$. AUT objective creativity scores ($p = .045$) and average Bloom score ($p < .001$) were statistically significant predictors.

For AUT objective creativity scores, a multiple regression analysis was applied to predict AUT objective creativity scores from openness, intelligence *Gf*, semantic fluency *Gr*, AQT objective creativity scores, curiosity, and Bloom average scores along with age, sex, and years of formal education as control variables. These variables significantly predicted AUT objective creativity scores, *F*(9,

Figure 5

Scatter Plots of the Relation of AQT Objective Creativity Scores With AUT Objective Creativity Scores, and AUT Objective Creativity Scores with Gf



Note. Circles represent average participants' scores; trendline represents correlations described above. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence.

105) = 2.067, $p = .039$, $R^2 = .156$. Only AQT objective creativity ($p = .024$) and *Gf* ($p = .042$) were statistically significant predictors. There were no problems of multicollinearity in either analysis. Similar to Study 1, VIF were examined to control for multicollinearity.

Serial Order Effect

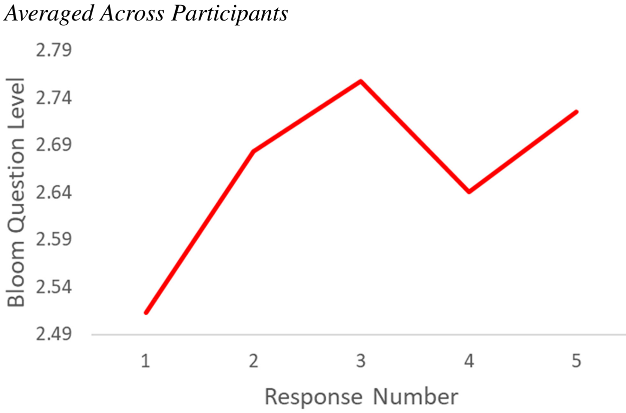
Although it was not included in the study preregistration, we conducted an exploratory ad hoc analysis on the effects of response order on average Bloom questions level across participants' responses on the AQT (Figure 6). Similar to Study 1, we conducted this analysis on participants' first five AQT responses. A one-way ANOVA revealed a significant main effect of response order, $F(4, 1476) = 2.633$, $p = .033$, $\eta_p^2 = .007$. Post hoc Bonferroni corrected paired-samples *t*-test analyses revealed significant differences between participants' serial order of responses, such that their first response had the lowest average Bloom level and their second and third responses had a significantly higher Bloom level (all $ps < .05$) than their first. While their fourth response was nonsignificant when compared to the first, participants' fifth response was also significantly higher than their first ($M_{diff} = 0.212$, 95% CI [0.031–0.393]), $p = .021$). These findings are in line with the serial order effect in creativity (Beaty & Silvia, 2012), where participant responses later in time are found to be more creative.

Discussion

Study 2 replicated and generalized almost all of the findings of Study 1, by finding similar relations between the AQT and AUT, after switching the items across the tasks used in Study 1. However, In the regression models predicting AUT objective creativity, *Gr* was significant in the Study 1, whereas *Gf* was significant in Study 2. Thus, Study 2 demonstrated that performance in the AQT is not object specific. In addition, unlike Study 1, we found a serial order effect of question asking (Beaty & Silvia, 2012): As the serial order of participants' questions increased after their first response, so did the Bloom level of the question.

Figure 6

Line Chart of Average Bloom Levels by Order of Response Averaged Across Participants



Note. All reponses average bloom levels except the fourth are significantly higher than the first response participants gave. See the online article for the color version of this figure.

General Discussion

Despite increased recent scientific interest (e.g., Gottlieb, 2021; Nelson, 2005; Rothe et al., 2018; Sasson & Kenett, 2023), there is still little research on how questions relate to creativity. Question asking has been shown to be highly important in children (Ronfard et al., 2018), and in higher-level creativity (Albergaria-Almeida, 2011)). Critically, question asking may play an important role in creativity during the problem finding stage of the creative process (Abdulla et al., 2020; Arreola & Reiter-Palmon, 2016). However, much is still unknown about specific types of questions asked and their relationship to creative thinking. In this study, we empirically and directly studied this issue.

While prior theory and empirical evidence have emphasized that divergent thinking tasks such as the AUT measure creativity (e.g., Acar & Runco, 2019), few so far besides the historical, no longer used, UQT have studied question asking skills and creativity (Torrance, 1970). Importantly, none of these scarce studies have evaluated the role of question asking in creativity using computational measures such as semantic distance and Bloom's taxonomy, to gain understanding on the role of asking more complex questions in creativity. This study addressed this research gap and capitalized on methodological advances in creativity scoring using semantic distance via the AQT.

The AQT examines how people ask questions about objects, and how this question asking ability relates to individual differences in creativity. We expected that questions higher on the Bloom taxonomy would be more semantically distant and subjectively rated as more creative by external judges (H1). The results supported these hypotheses, showing how higher-level questions were positively correlated with both objective (semantic distance) and subjective (externally rated) creativity scores. We also expected frequencies of high and low Bloom level questions to correlate differently with creativity (H2). Indeed, individual Bloom questions levels were differently related to their creativity scores: Whereas higher frequencies of lower-level questions such as remembering and understanding were negatively related to AQT creativity scores, higher-level questions such as analyzing and evaluating were positively related to AQT creativity scores.

Additionally, in both studies, average Bloom level scores were normally distributed, which strengthened the assumption that higher Bloom levels refer to more cognitively complex questions, while lower levels represent lower complexity, with the average participant gravitating toward average Bloom level (H3). Nevertheless, further research is required to validate these findings.

The results also showed that participants differed in semantic distance scores, generating more distant responses to the AQT compared to the AUT in both studies. Furthermore, as hypothesized, we found a significant, but modest correlation between the AQT and AUT (H4), thus coupling these two capacities—generating uses and asking questions—together in relation to creativity yet at the same time setting them apart as different. Taken together, these findings indicate that the AUT and AQT measure different aspects of divergent thinking. Given the likely role of question asking in problem finding (Reiter-Palmon & Robinson, 2009), our findings highlighted the significance of assessing question asking in creativity research.

One interpretation of these findings is that higher-level questions, which require more complicated synthesis of information, require participants to think divergently, and to link remote ideas in a creative way. Perhaps, some of the same skills that allow us to be creative in the AUT, enable us to phrase more complex and creative questions. Surprisingly, we found in Study 1 that fluency in the AQT and AUT are negatively correlated with AQT objective and subjective creativity, as well as Bloom level scores. This finding is not confounded by response length, given our use of the MAD method (which controls for variation in response elaboration; Yu et al., 2023). In the AUT, being more fluent facilitates the generation of more creative ideas (Bai et al., 2021; Beaty & Silvia, 2012). However, in Study 2, we did find a negative correlation between AUT fluency and AUT MAD scores. Perhaps, in the AQT, being more fluent in asking questions may result in producing simpler and shorter responses. These responses are then more likely to be less complex and lower-level questions, which we have found to be less creative. This is in line with the tradeoff model proposed by Forthmann et al. (2020), which builds upon work done in Guilford's lab which proposed that there are true individual differences in average creativity that are negatively correlated with fluency. Guilford has even stated that "if a person spends his or her time producing a lot of low-quality responses, he or she cannot produce so many good ones" (Guilford, 1968, p. 104). This has also been observed in Warne et al. (2021) who found that creativity fluency scores were negatively correlated with originality, suggesting that there is a tradeoff, where generating many responses, tends to reduce scores on originality aspects of divergent thinking (see also Beaty et al., 2023). Reiter-Palmon and Arreola (2015) found similar effects when participants were asked to generate multiple solutions to a real-world problem compared to one solution. When generating multiple ideas, solutions were less complex and of lower creativity compared to generating one solution.

In contrast, we found an effect of response order in Study 2, where average Bloom levels significantly rose with AQT response order, such that besides their fourth response, participant's average Bloom question level kept increasing until their fifth response, although this effect was not observed in Study 1. This is a surprising finding as it seems to contradict the negative effect of AQT fluency on Bloom question level. Perhaps although fluency as a whole lowered AQT Bloom level when participants list many possible questions (some even listing more than 20 questions): In the context of the early (first to fifth) responses, participants firstly exhaust simpler, less complex questions and slowly construct more complex ones until their fifth response after which they dip again. Although this analysis was done ad hoc, owing to the nature of this study as exploratory it provides some direction for future research on response order and fluency in the AQT. This finding also replicated work by Beaty and Silvia (2012) who found a similar relationship using the AUT (see also Bai et al., 2021 for similar findings in children).

Our hypotheses regarding the AQT and additional assessed variables (H5) were based on the relation between divergent thinking and intelligence (Batey et al., 2009; Lee & Therriault, 2013; Silvia, 2008, 2015), openness to experience (Batey & Furnham, 2006), and curiosity (Celik et al., 2016; Vidler & Karan, 1975) that performance in the AQT will also be related to *Gf*, *Gr*, openness, and curiosity scores as will performance on the AUT. In Study 1 we found that besides modest correlations with *Gf* and semantic fluency *Gr* with AQT subjective creativity, AQT objective creativity did not correlate with *Gf*, curiosity, openness to experience, or semantic fluency *Gr*. This may be in part due to the MAD methods' tendency to capture idea novelty and not usefulness (Yu et al., 2023). Further regression analyses found that for predicting AUT objective creativity only *Gr* was a significant predictor while *Gf*, curiosity and openness were nonsignificant for predicting AQT objective creativity. Additionally, when examining the frequency of individual Bloom levels of participants' responses, in Study 1 Bloom Level 3 was modestly positively related to semantic fluency *Gr*, while in Study 2 Bloom Level 1 was negatively correlated with curiosity and openness to experience. Furthermore, Bloom Level 2 was negatively correlated with curiosity and Bloom Level 3 was nonsignificant across all variables. Bloom Level 4 was significantly positively related to Semantic Fluency *Gr*, while interestingly Bloom Level 5 was also modestly negatively related to curiosity. Bloom Level 6 was nonsignificant across all variables.

These correlations, although modest, provide some strength to the argument that higher Bloom levels in the taxonomy refer to more cognitively complex questions, especially in line with literature findings that higher semantic knowledge can be represented as a more complex semantic memory network (Goñi et al., 2011) and semantic fluency's relation to executive functions (Ardila et al., 2006).

In Study 2 we found similar modest correlations with *Gf* and *Gr* with AQT subjective creativity scores, which is in line with previous research that has linked creativity with *Gf* and *Gr* (Beaty & Silvia, 2012; Silvia, 2015; Silvia et al., 2013). Further regression analyses found that for predicting AUT objective creativity only *Gf* was a significant predictor while *Gr*, curiosity, and openness were nonsignificant for predicting AQT objective creativity. Similar to Study 1, we did not find any significant correlations with AQT objective creativity scores and *Gf*, curiosity, openness to experience, and *Gr*. These findings provide further support to the notion that although they correlate highly, there is still some difference between subjective and objective AQT creativity scores.

Study Limitations and Implications

There are some limitations concerning the results of this study. First, is the usage of MAD semantic distance scores as a quantitative measure of creativity. According to Yu et al. (2023), MAD scores most likely capture mainly elaboration and idea novelty or originality, one of the many facets of creative performance. Although MAD scores do not explicitly address the appropriateness (or usefulness) of the response, Heinen and Johnson (2018) have shown that by simply emphasizing that the goal of a task is to "be creative," as we have in our study, semantic distance becomes a remarkably robust measure for assessing creativity that is novel and appropriate. When this is not emphasized, semantic distance values index novelty more closely than creativity. Beaty and Johnson (2021) have found that when novelty is emphasized over appropriateness, the correlation between human ratings and semantic distance increases. A second limitation is the relatively low reliability of subjectively rated creativity by external raters in both studies. Although considered moderate by Koo and Li's (2016) interrater reliability standards, it impacts some of the study's external validity conclusions. As such, further research is needed to improve the way of rating creativity of AQT responses, especially via laypersons (Hass et al., 2018; Wang et al., 2023). Perhaps more effort is needed in training raters to increase the reliability and improve the AQT's validity as a tool used to examine creativity, similar to Amabile's (1982) consensual assessment technique or utilization of quasi-expert raters (Kaufman & Baer, 2012; Kaufman et al., 2013).

Additionally, it is important to note that subjective scoring of responses according to the Bloom taxonomy may suffer from the same limitations as subjectively rated creativity, such as inconsistent rater agreement, rater fatigue and high costs. But, whereas objective measures of semantic distance correlate highly and overlap with subjectively rated creativity, the Bloom taxonomy examines additional facets such as complexity and cognitive hierarchy in question asking and creativity, which are not accounted for in semantic distance. Automated, computational Bloom taxonomy scoring may thus be beneficial in overcoming some of these limitations and further research should explore this possibility. Lastly, although we examined both objective semantic distance creativity scores with externally rated subjective creativity scores, our task was more a controlled, lab study. Thus, further tests of real-life creativity problems are needed to extend the external validity of our findings.

Despite these limitations, the results suggest several theoretical and practical implications. Firstly, we found a strong relation between question-level and objective and subjective creativity. This finding holds many practical implications for building better creativity tests. Likewise, the implementation of Bloom's taxonomy as an additional measure of creative ability that is both related to and different from divergent thinking according to our findings, can help further understanding of creative ability, by bringing attention to question complexity in the creative process. The findings highlighted in this study may also demonstrate the potential for utilizing the Bloom taxonomy as a creativity scale in additional areas, beyond even question asking.

Furthermore, research with elementary school and college-aged students has shown that students can quickly be taught how to ask higher-level questions, and that this leads to improvements in learning and reading comprehension (King, 1990, 1992; Rosenshine et al., 1996). This may indicate the potential benefits of further research on the impact of question level manipulations and interventions, which could pave the way for improvements in learning abilities.

Finally, the differences we find between the AUT and AQT opens new lines of research on the relation between fluency and creative thought. This may indicate that question asking is more complex than the relation examined here. Perhaps question asking and the AQT specifically are related to problem finding processes, which can be examined using the questions asked during attempts to solve ill-defined problems, as was tested by Wigert et al. (2022). Further work linking question asking and high-level cognition is needed, such as extending the current findings by examining additional measures of creativity and analyzing them according to Bloom's taxonomy. Future research is also needed to examine complexity of question asking in a more quantitative way than conducted in this study. Nevertheless, our study sheds unique light on the role of complex questions and creativity, highlighting the significance of heightened question asking ability. Given that question asking can lead to problem finding, that in turn leads to problem solving, question asking is a critical cognitive capacity.

Conclusions

In sum, our research is an important first step toward integrating two lines of research, creativity and Bloom taxonomy rated questions. The current work suggests that the AQT coupled with Bloom's taxonomy is an important tool for understanding and testing creative and question asking abilities, and that the AQT is more than just a divergent thinking task. It also highlights the possible advantages of testing creative thinking with additional measures besides the AUT. As Cortes et al. (2019) claims, the AUT cannot capture creativity as a whole and research should seek to utilize a battery of tasks. Importantly, this study allows us to move forward in finding more complex ways to ask: What is the role of question asking—and especially more complex questions—in creative thinking?

References

Abdulla, A. M., Paek, S. H., Cramond, B., & Runco, M. A. (2020). Problem finding and creativity: A meta-analytic review. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 3–14. https://doi.org/10.1037/aca0000194

Abraham, A., & Bubic, A. (2015). Semantic memory as the root of imagination. *Frontiers in Psychology*, *6*, Article 325. https://doi.org/10.3389/fpsyg.2015.00325

Acar, S., Berthiaume, K., & Johnson, R. (2023). What kind of questions do creative people ask? *Journal of Creativity*, *33*(3), Article 100062. https://doi.org/10.1016/j.yjoc.2023.100062

Acar, S., & Runco, M. A. (2019). Divergent thinking: New methods, recent research, and extended theory. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 153–158. https://doi.org/10.1037/aca0000231

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*, *103*(3), 152–153. https://doi.org/10.3163/1536-5050.103.3.010

Alabbasi, A. M. A., Acar, S., & Reiter-Palmon, R. (2023). *Problem finding and divergent thinking: A multivariate meta-analysis*. PsyArxiv. https://doi.org/10.31219/osf.io/u9pn2

Albergaria-Almeida, P. (2011). Critical thinking, questioning and creativity as components of intelligence. *Procedia—Social and Behavioral Sciences*, *30*, 357–362. https://doi.org/10.1016/j.sbspro.2011.10.070

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*(5), 997–1013. https://doi.org/10.1037/0022-3514.43.5.997

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of semantic verbal fluency (ANIMALS). *International Journal of Psychology*, *41*(5), 324–332. https://doi.org/10.1080/00207590500345542

Arreola, N. J., & Reiter-Palmon, R. (2016). The effect of problem construction creativity on solution creativity across multiple everyday problems. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(3), 287–295. https://doi.org/10.1037/a0040389

Avitia, M. J., & Kaufman, J. C. (2014). Beyond *g* and *c*: The relationship of rated creativity to long-term storage and retrieval (*Glr*). *Psychology of Aesthetics, Creativity, and the Arts*, *8*(3), 293–302. https://doi.org/10.1037/a0036772

Baer, J. (2012). Domain specificity and the limits of creativity theory. *The Journal of Creative Behavior*, *46*(1), 16–29. https://doi.org/10.1002/jocb.002

Bai, H., Mulder, H., Moerbeek, M., Kroesbergen, E. H., & Leseman, P. P. M. (2021). Divergent thinking in four-year-old children: An analysis of thinking processes in performing the alternative uses task. *Thinking Skills and Creativity*, *40*, Article 100814. https://doi.org/10.1016/j.tsc.2021.100814

Baloche, L. (1994). Breaking down the walls. *The Social Studies*, *85*(1), 25–30. https://doi.org/10.1080/00377996.1994.10118776

Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2009). Intelligence and personality as predictors of divergent thinking: The role of general, fluid and crystallized intelligence. *Thinking Skills and Creativity*, *4*(1), 60–69. https://doi.org/10.1016/j.tsc.2009.01.002

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. https://doi.org/10.3200/MONO.132.4.355-430

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780. https://doi.org/10.3758/s13428-020-01453-w

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780. https://doi.org/10.3758/s13428-020-01453-w

Beaty, R. E., Johnson, D. R., Zeitlen, D. C., & Forthmann, B. (2022). Semantic distance and the alternate uses task: Recommendations for reliable automated assessment of originality. *Creativity Research Journal*, *34*(3), 245–260. https://doi.org/10.1080/10400419.2022.2025720

Beaty, R. E., & Kenett, Y. N. (2023). Associative thinking at the core of creativity. *Trends in Cognitive Sciences*, *27*(7), 671–683. https://doi.org/10.1016/j.tics.2023.04.004

Beaty, R. E., Kenett, Y. N., Hass, R. W., & Schacter, D. L. (2023). Semantic memory and creativity: The costs and benefits of semantic memory structure in generating original ideas. *Thinking & Reasoning*, *29*(2), 305–339. https://doi.org/10.1080/13546783.2022.2076742

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative over time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 309–319. https://doi.org/10.1037/a0029171

Benedek, M., Beaty, R. E., Schacter, D., & Kenett, Y. N. (2023). The role of memory in creative ideation. *Nature Reviews Psychology*, *2*(4), 246–257. https://doi.org/10.1038/s44159-023-00158-z

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1956). *Taxonomy of educational objectives: The classification of educational goals*. David McKay Company.

Burkhart, R. C. (1961). The interrelationship of separate criteria for creativity in art and student teaching to four personality factors. *Studies in Art Education*, *3*(1), 18–38. https://doi.org/10.2307/1319526

Carroll, J. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. https://doi.org/10.1037/h0046743

Cattell, R. B., & Cattell, A. K. S. (1961/2008). *Measuring intelligence with the Culture Fair Tests*. Hogrefe.

Celik, P., Storme, M., Davila, A., & Myszkowski, N. (2016). Work-related curiosity positively predicts worker innovation. *Journal of Management Development*, *35*(9), 1184–1194. https://doi.org/10.1108/JMD-01-2016-0013

Collins, R. P., Litman, J. A., & Spielberger, C. D. (2004). The measurement of perceptual curiosity. *Personality and Individual Differences*, *36*(5), 1127–1141. https://doi.org/10.1016/S0191-8869(03)00205-8

Conner, T. S., & Silvia, P. J. (2015). Creative days: A daily diary study of emotion, personality, and everyday creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 463–470. https://doi.org/10.1037/aca0000022

Cortes, R. A., Weinberger, A. B., Daker, R. J., & Green, A. E. (2019). Re-examining prominent measures of divergent and convergent creativity. *Current Opinion in Behavioral Sciences*, *27*, 90–93. https://doi.org/10.1016/j.cobeha.2018.09.017

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5–13. https://doi.org/10.1037/1040-3590.4.1.5

Curto, J. D., & Pinto, J. C. (2011). The corrected VIF (CVIF). *Journal of Applied Statistics*, *38*(7), 1499–1507. https://doi.org/10.1080/02664763.2010.505956

Dumas, D., Organisciak, P., & Doherty, M. (2021). Measuring divergent thinking originality with human raters and text-mining models: A psychometric comparison of methods. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(4), 645–663. https://doi.org/10.1037/aca0000319

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Forehand, M. (2010). Bloom's taxonomy. *Emerging Perspectives on Learning, Teaching, and Technology*, *41*(4), 47–56. https://www.d41.org/cms/lib/il01904672/centricity/domain/422/bloomstaxonomy.pdf

Forthmann, B., Holling, H., Zandi, N., Gerwig, A., Çelik, P., Storme, M., & Lubart, T. (2017). Missing creativity: The effect of cognitive workload on rater (dis-)agreement in subjective divergent-thinking scores. *Thinking Skills and Creativity*, *23*, 129–139. https://doi.org/10.1016/j.tsc.2016.12.005

Forthmann, B., Szardenings, C., & Holling, H. (2020). Understanding the confounding effect of fluency in divergent thinking scores: Revisiting average scores to quantify artifactual correlation. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 94–112. https://doi.org/10.1037/aca0000196

Getzels, J. W. (1979). Problem finding: A theoretical note. *Cognitive Science*, *3*(2), 167–172. https://doi.org/10.1207/s15516709cog0302_4

Goh, T. T., Mohamed, H., Jamaludin, N. A., Ismail, M. N., & Chua, H. S. (2020). Questions classification according to Bloom's taxonomy using universal dependency and Word Net. *Test Engineering and Management*, *82*, 4374–4385. https://www.testmagzine.biz/index.php/testmagzine/article/view/1525/1380

Goñi, J., Arrondo, G., Sepulcre, J., Martincorena, I., Vélez de Mendizábal, N., Corominas-Murtra, B., Bejarano, B., Ardanza-Trevijano, S., Peraita, H., Wall, D. P., & Villoslada, P. (2011). The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cognitive Processing*, *12*(2), 183–196. https://doi.org/10.1007/s10339-010-0372-x

Gottlieb, J. (2021). The effort of asking good questions. *Nature Human Behaviour*, *5*(7), 823–824. https://doi.org/10.1038/s41562-021-01132-6

Grajzel, K., Dumas, D., & Acar, S. (2022). Are the verbal TTCT forms actually interchangeable? *The Journal of Creative Behavior*, *56*(4), 629–646. https://doi.org/10.1002/jocb.556

Granello, D. H. (2001). Promoting cognitive complexity in graduate written work: Using Bloom's taxonomy as a pedagogical tool to improve literature reviews. *Counselor Education and Supervision*, *40*(4), 292–307. https://doi.org/10.1002/j.1556-6978.2001.tb01261.x

Grebin, N., Grabovska, S., Karkovska, R., & Vovk, A. (2020). Applying Benjamin Bloom's taxonomy ideas in adult learning. *Journal of Education Culture and Society*, *11*(1), 61–72. https://doi.org/10.15503/jecs2020.1.61.72

Gross, M. E., Zedelius, C. M., & Schooler, J. N. (2020). Cultivating an understanding of curiosity as a seed for creativity. *Current Opinion in Behavioral Sciences*, *35*, 77–82. https://doi.org/10.1016/j.cobeha.2020.07.015

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.

Guilford, J. P. (1968). *Intelligence, creativity, and their educational implications*. Robert R. Knapp.

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006–1033. https://doi.org/10.1177/1745691619861372

Hass, R. W. (2017a). Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory & Cognition*, *45*(2), 233–244. https://doi.org/10.3758/s13421-016-0659-y

Hass, R. W. (2017b). Semantic search during divergent thinking. *Cognition*, *166*, 344–357. https://doi.org/10.1016/j.cognition.2017.05.039

Hass, R. W., Rivera, M., & Silvia, P. J. (2018). On the dependability and feasibility of layperson ratings of divergent thinking. *Frontiers in Psychology*, *9*, Article 1343. https://doi.org/10.3389/fpsyg.2018.01343

Heinen, D. J. P., & Johnson, D. R. (2018). Semantic distance: An automated measure of creativity that is novel and appropriate. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(2), 144–156. https://doi.org/10.1037/aca0000125

Kashdan, T. B., Disabato, D. J., Goodman, F. R., & McKnight, P. E. (2020). The five- dimensional curiosity scale revised (5DCR): Briefer subscales while separating overt and covert social curiosity. *Personality and Individual Differences*, *157*, Article 109836. https://doi.org/10.1016/j.paid.2020.109836

Kaufman, J. C. (2019). Self-assessments of creativity: Not ideal, but better than you think. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 187–192. https://doi.org/10.1037/aca0000217

Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal*, *24*(1), 83–91. https://doi.org/10.1080/10400419.2012.649237

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*(4), 332–340. https://doi.org/10.1037/a0034809

Kenett, Y. N. (2018). Going the extra creative mile: The role of semantic distance in creativity—Theory, research, and measurement. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge handbook of the neuroscience of creativity* (pp. 233–248). Cambridge University Press. https://doi.org/10.1017/9781316556238.014

Kenett, Y. N. (2019). What can quantitative measures of semantic distance tell us about creativity? *Current Opinion in Behavioral Sciences*, *27*, 11–16. https://doi.org/10.1016/j.cobeha.2018.08.010

Kenett, Y. N., & Faust, M. (2019). A semantic network cartography of the creative mind. *Trends in Cognitive Sciences*, *23*(4), 271–274. https://doi.org/10.1016/j.tics.2019.01.007

Kenett, Y. N., Humphries, S., & Chatterjee, A. (2023). A thirst for knowledge: Grounding creativity, curiosity, and aesthetic experience in memory and reward neural systems. *Creativity Research Journal*, *35*(3), 412–426. https://doi.org/10.1080/10400419.2023.2165748

Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior*, *42*(2), 106–130. https://doi.org/10.1002/j.2162-6057.2008.tb01290.x

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*(4), 664–687. https://doi.org/10.3102/00028312027004664

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal*, *29*(2), 303–323. https://doi.org/10.3102/00028312029002303

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Koutstaal, W., Kedrick, K., & Gonzalez-Brito, J. (2022). Capturing, clarifying, and consolidating the curiosity-creativity connection. *Scientific Reports*, *12*(1), Article 15300. https://doi.org/10.1038/s41598-022-19694-4

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*(5), 306–320. https://doi.org/10.1016/j.intell.2013.04.008

Litman, J. A. (2008). Interest and deprivation dimensions of epistemic curiosity. *Personality and Individual Differences*, *44*(7), 1585–1595. https://doi.org/10.1016/j.paid.2008.01.014

Litman, J. A., & Jimerson, T. L. (2004). The measurement of curiosity as a feeling of deprivation. *Journal of Personality Assessment*, *82*(2), 147–157. https://doi.org/10.1207/s15327752jpa8202_3

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment*, *80*(1), 75–86. https://doi.org/10.1207/S15327752JPA8001_16

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, *52*(6), 1258–1265. https://doi.org/10.1037/0022-3514.52.6.1258

McCrae, R. R., & Greenberg, D. M. (2014). Openness to experience. In D. K. Simonton (Ed.), *The Wiley handbook of genius* (pp. 222–243). Wiley Blackwell. https://doi.org/10.1002/9781118367377.ch12

Mumford, M. D., Mobley, M. I., Uhlman, C. E., Reiter-Palmon, R., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*(2), 91–122. https://doi.org/10.1080/10400419109534380

Mumford, M. D., Reiter-Palmon, R., & Redmond, M. R. (1994). Problem construction and cognition: Applying problem representations in ill-defined domains. In M. A. Runco (Ed.), *Problem finding, problem solving, and creativity* (pp. 3–39). Ablex. https://psycnet.apa.org/record/1994-98186-001.

Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4), 979–999. https://doi.org/10.1037/0033-295X.112.4.979

Okuda, S. M., Runco, M. A., & Berger, D. E. (1991). Creativity and the finding and solving of real-world problems. *Journal of Psychoeducational Assessment*, *9*(1), 45–53. https://doi.org/10.1177/073428299100900104

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. Fiest, R. Reiter-Palmon, & J. C. Kaufman (Eds.), *Cambridge handbook of creativity and personality research* (pp. 9–27). Cambridge University Press. https://doi.org/10.1017/9781316228036.002.

Oliver, D., Dobele, T., Greber, M., & Roberts, T. S. (2004, January 18–22). *This course has a Bloom rating of 3.9* [Symposium]. IFAC Symposium on Advances in Control Education, Dunedin, New Zealand.

Omar, N., Haris, S. S., Hassan, R., Arshad, H., Rahmat, M., Zainal, N. F., & Zulkifli, R. (2012). Automated analysis of exam questions according to Bloom's taxonomy. *Procedia—Social and Behavioral Sciences*, *59*, 297–303. https://doi.org/10.1016/j.sbspro.2012.09.278

Organisciak, P., Acar, S., Dumas, D., & Berthiaume, K. (2023). Beyond semantic distance: Automated scoring of divergent thinking greatly improves with large language models. *Thinking Skills and Creativity*, *49*, Article 101356. https://doi.org/10.1016/j.tsc.2023.101356

Plack, M. M., Driscoll, M., Marquez, M., Cuppernull, L., Maring, J., & Greenberg, L. (2007). Assessing reflective writing on a pediatric clerkship by using a modified Bloom's taxonomy. *Ambulatory Pediatrics*, *7*(4), 285–291. https://doi.org/10.1016/j.ambp.2007.04.006

Puryear, J. S., Kettler, T., & Rinn, A. N. (2017). Relationships of personality to differential conceptions of creativity: A systematic review. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(1), 59–68. https://doi.org/10.1037/aca0000079

Reiter-Palmon, R., & Arreola, N. J. (2015). Does generating multiple ideas lead to increased creativity? A comparison of generating one idea vs. many. *Creativity Research Journal*, *27*(4), 369–374. https://doi.org/10.1080/10400419.2015.1087274

Reiter-Palmon, R., Mumford, M. D., O'Connor Boes, J., & Runco, M. A. (1997). Problem construction and creativity: The role of ability, cue consistency, and active processing. *Creativity Research Journal*, *10*(1), 9–23. https://doi.org/10.1207/s15326934crj1001_2

Reiter-Palmon, R., Mumford, M. D., & Threlfall, K. V. (1998). Solving everyday problems creatively: The role of problem construction and personality type. *Creativity Research Journal*, *11*(3), 187–197. https://doi.org/10.1207/s15326934crj1103_1

Reiter-Palmon, R., & Robinson, E. (2009). Problem identification and construction: What do we know, what is the future? *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 43–47. https://doi.org/10.1037/a0014629

Roberts, N. (1976). Further verification of Bloom's taxonomy. *The Journal of Experimental Education*, *45*(1), 16–19. https://doi.org/10.1080/00220973.1976.11011556

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. *Developmental Review*, *49*, 101–120. https://doi.org/10.1016/j.dr.2018.05.002

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, *66*(2), 181–221. https://doi.org/10.3102/00346543066002181

Rothe, A., Lake, B. M., & Gureckis, T. M. (2018). Do people ask good questions? *Computational Brain & Behavior*, *1*(1), 69–89. https://doi.org/10.1007/s42113-018-0005-5

Runco, M. A. (1986). Divergent thinking and creative performance in gifted and nongifted children. *Educational and Psychological Measurement*, *46*(2), 375–384. https://doi.org/10.1177/001316448604600211

Runco, M. A. (1992). Children's divergent thinking and creative ideation. *Developmental Review*, *12*(3), 233–264. https://doi.org/10.1016/0273-2297(92)90010-Y

Runco, M. A. (2010). Divergent thinking, creativity, and ideation. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 413–446). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.026

Runco, M. A., Illies, J. J., & Reiter-Palmon, R. (2005). Explicit instructions to be creative and original: A comparison of strategies and criteria as targets with three types of divergent thinking tests. *Korean Journal of Thinking & Problem Solving*, *15*(1), 5–15.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92–96. https://doi.org/10.1080/10400419.2012.650092

Runco, M. A., & Mraz, W. (1992). Scoring divergent thinking tests using total ideational output and a creativity index. *Educational and Psychological Measurement*, *52*(1), 213–221. https://doi.org/10.1177/001316449205200126

Runco, M. A., & Nemiro, J. (1994). Problem finding, creativity, and giftedness. *Roeper Review*, *16*(4), 235–241. https://doi.org/10.1080/02783199409553588

Said-Metwaly, S., Fernández-Castilla, B., Kyndt, E., & Van den Noortgate, W. (2020). Testing conditions and creative performance: Meta-analyses of the impact of time limits and instructions. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 15–38. https://doi.org/10.1037/aca0000244

Said-Metwaly, S., Taylor, C. L., Camarda, A., & Barbot, B. (2022). Divergent thinking and creative achievement—How strong is the link? An updated meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000507

Sasson, G., & Kenett, Y. N. (2023). A mirror to human question asking: Analyzing the akinator online question game. *Big Data and Cognitive Computing*, *7*(1), Article 26. https://doi.org/10.3390/bdcc7010026

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*, *9*(6), 523–538. https://doi.org/10.1002/acp.2350090605

Schutte, N. S., & Malouff, J. M. (2020). A meta-analysis of the relationship between curiosity and creativity. *The Journal of Creative Behavior*, *54*(4), 940–947. https://doi.org/10.1002/jocb.421

Silvia, P. J. (2008). Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). *Creativity Research Journal*, *20*(1), 34–39. https://doi.org/10.1080/10400410701841807

Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychology Review*, *27*(4), 599–606. https://doi.org/10.1007/s10648-015-9299-1

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence*, *41*(5), 328–340. https://doi.org/10.1016/j.intell.2013.05.004

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & 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. https://doi.org/10.1037/1931-3896.2.2.68

Smith, I. L. (1970). IQ, creativity, and the taxonomy of educational objectives. *The Journal of Experimental Education*, *38*(4), 58–60. https://doi.org/10.1080/00220973.1970.11011215

Sternberg, R. J., & Williams, W. M. (1996). *How to develop student creativity*. ASCD Press.

Thurstone, L. L. (1938). *Primary mental abilities*. University of Chicago Press.

Torrance, E. P. (1966). *Torrance tests of creative thinking, directions manual and scoring guide*. Personnel Press.

Torrance, E. P. (1970). Group size and question performance of preprimary children. *The Journal of Psychology*, *74*(1), 71–75. https://doi.org/10.1080/00223980.1970.10545279

Van Der Meij, H. (1994). Student questioning: A componential analysis. *Learning and Individual Differences*, *6*(2), 137–161. https://doi.org/10.1016/1041-6080(94)90007-8

Vidler, D. C., & Karan, V. E. (1975). A study of curiosity, divergent thinking, and test-anxiety. *The Journal of Psychology*, *90*(2), 237–243. https://doi.org/10.1080/00223980.1975.9915781

Volle, E. (2018). Associative and controlled cognition in divergent thinking: Theoretical, experimental, neuroimaging evidence, and new directions. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge handbook of the neuroscience of creativity* (pp. 333–360). Cambridge University Press. https://doi.org/10.1017/9781316556238.020

Wang, K., Dong, B., & Ma, J. (2023) Testing computational assessment of idea novelty in crowdsourcing. *Creativity Research Journal*. Advance online publication. https://doi.org/10.1080/10400419.2023.2187544

Warne, R. T., Golightly, S., & Black, M. (2021). Factor structure of intelligence and divergent thinking subtests: A registered report. *PLoS ONE*, *16*(5), Article e0251268. https://doi.org/10.1371/journal.pone.0251268

Wigert, B. G., Murugavel, V. R., & Reiter-Palmon, R. (2022). The utility of divergent and convergent thinking in the problem construction processes during creative problem-solving. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000513

Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z. X., Chiu, C. Y., & Stark, S. E. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, *96*(1), 29–45. https://doi.org/10.1080/00223891.2013.806328

Yang, W., Green, A. E., Chen, Q., Kenett, Y. N., Sun, J., Wei, D., & Qiu, J. (2022). Creative problem solving in knowledge-rich contexts. *Trends in Cognitive Sciences*, *26*(10), 849–859. https://doi.org/10.1016/j.tics.2022.06.012

Yu, Y., Beaty, R., Forthmann, B., Beeman, M., Cruz, J. H., & Johnson, D. R. (2023). A mad method to assess idea novelty: Improving validity of automatic scoring using maximum associative distance (MAD). *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.31234/osf.io/vgxpk

Zheng, A. Y., Lawhorn, J. K., Lumley, T., & Freeman, S. (2008). Assessment. Application of Bloom's taxonomy debunks the "MCAT myth". *Science*, *319*(5862), 414–415. https://doi.org/10.1126/science.1147852

(*Appendices follow*)

Appendix A

Subjective Creativity Rating Instructions (Example for Pencil)

Please read through the following questions asked about a pencil, which were obtained from different participants.

For each question, you are asked to assign a numerical rating indicating how much you think this question is creative.

The rating scale is between 1 and 5 based on your own understanding of what a creative idea is, with 1 being the least creative and 5 being the most creative.

It is very important that you use the full 1–5 scale, and not assign almost all of the responses the same rating.

Appendix B

Bloom Taxonomy Rating Instructions (Example for Pencil)

"In the following experiment you will see questions asked about a pencil.

People ask all sorts of questions and we are interested in assigning the questions that were asked about a pencil into their different types based on a six-level classification taxonomy."

Here are some types of questions that people usually ask, please read them thoroughly and try to understand the different question levels:

1. Remembering—These types of questions address terms, facts and details without necessarily understanding the concept. For example, What is it…? Where is it…? What color is it…?
2. Understanding—These questions test the ability to summarize and describe in our own words without necessarily relating it to anything. For example, What is meant by…? How would you explain…? What can you say about…?
3. Applying—These questions involve applying or transferring learning to our own life or to a context different than one in which we learned. For example, How would you use…? What facts would you select to show…? What would happen if…?
4. Analyzing—These questions are about breaking material into parts, describing patterns and relationships among parts, to subdivide information and to show how it is put together. For example, What is the relationship between…? Why do you think…? What is the function of…?
5. Evaluating—These questions involve developing opinions and making value decisions about issues based on specific criteria. For example, What is your opinion of…? Is it better/worse than…? How would you test…? What way would you design…?
6. Creating—These questions involve creating something new by using a combination of ideas from different sources to form a new whole. For example, What changes would you make…? What could be combined to improve…?

For example:

What is a pencil? would be Level 1 creating.
What does the name pencil mean? would be Level 2—understanding
Can you use it eat something? would be Level 3—applying
What is the purpose of the wood in the pencil? would be Level 4—analyzing
Could this pencil be colored red? would be Level 5—evaluating
Is a pencil better than a pen? would be Level 6—creating.

Please evaluate the following questions asked about a pencil according to the different question types. After selecting an answer, the next question will automatically appear and there is no option to go back, so please click an answer only after you are sure that is what you would like to select."

(*Appendices continue*)

Appendix C

Correlation Table for Study 1

Variables		1	2	3	4	5	6	7	8	9	10	11
1. AUT fluency	Pearson correlation	—										
	N	109										
2. AUT mad	Pearson correlation	.122	—									
	Sig. (two-tailed)	.206										
	N	109	109									
3. AQT fluency	Pearson correlation	.445**	.004	—								
	Sig. (two-tailed)	.000	.969									
	N	109	109	109								
4. AQT subjective creativity	Pearson correlation	−.035	.318**	−.343**	—							
	Sig. (two-tailed)	.717	.001	.000								
	N	108	108	108	108							
5. AQT mad	Pearson correlation	.025	.247**	−.333**	.543**	—						
	Sig. (two-tailed)	.799	.010	.000	.000							
	N	109	109	109	108	109						
6. AQT bloom level	Pearson correlation	−.064	.188	.249**	.552**	.510**	—					
	Sig. (two-tailed)	.513	.51	.019	.000	.000						
	N	108	108	108	108	108	108					
7. *Gf*	Pearson correlation	.070	.249**	−.027	.238*	.082	−.002	—				
	Sig. (two-tailed)	.471	.009	.781	.013	.394	.984					
	N	109	109	109	108	109	108	109				
8. Curiosity score	Pearson correlation	−.215*	−.013	−.011	.016	.060	−.063	−.036	—			
	Sig. (two-tailed)	.025	.894	.911	.873	.538	.515	.710				
	N	109	109	109	108	109	108	109	109			
9. Openness	Pearson correlation	−.142	−.001	.058	.064	−.006	−.056	.026	.675**	—		
	Sig. (two-tailed)	.139	.990	.546	.511	.947	.565	.790	.000			
	N	109	109	109	108	109	108	109	109	109		
10. Semantic fluency *Gr*	Pearson correlation	.279**	.334**	.240*	.277**	0.094	0.002	.480**	−0.144	−0.008	.364**	—
	Sig. (two-tailed)	.003	.000	.012	.004	.332	.981	.000	.135	.937	.000	
	N	109	109	109	108	109	108	109	109	109	109	109

Note. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability; Sig. = significance.
$^{*}p < .05$. $^{**}p < .01$.

(*Appendices continue*)

Appendix D

Correlation Table for Study 2

Variables		1	2	3	4	5	6	7	8	9	10
1. AUT fluency	Pearson correlation	—									
	N	113									
2. AUT mad	Pearson correlation	−.198*	—								
	Sig. (two-tailed)	.018									
	N	113	113								
3. AQT fluency	Pearson correlation	−.090	.126	—							
	Sig. (two-tailed)	.172	.092								
	N	113	113	113							
4. AQT subjective creativity	Pearson correlation	.059	.149	−.086	—						
	Sig. (two-tailed)	.268	.127	.447							
	N	113	113	113	113						
5. AQT mad	Pearson correlation	.050	.215*	−.132	.362**	—					
	Sig. (two-tailed)	.300	.023	.081	.000						
	N	113	113	113	113	113					
6. AQT bloom level	Pearson correlation	−.015	.086	−.158	.413**	.470**	—				
	Sig. (two-tailed)	.876	.366	.097	.000	.016					
	N	113	113	113	113	113	113				
7. *Gf*	Pearson correlation	−.027	−.015	.082	.311**	.098	.120	—			
	Sig. (two-tailed)	.778	.874	.389	.000	.305	.207				
	N	113	113	113	113	113	113	113			
8. Curiosity score	Pearson correlation	−.075	.147	−.006	.103	.196*	.095	−.089	—		
	Sig. (two-tailed)	.216	.060	.474	.139	.037	.261	.174			
	N	113	113	113	113	113	113	113	113		
9. Openness	Pearson correlation	−.141	.099	.146	.178	.156	.077	.091	.537**	—	
	Sig. (two-tailed)	.068	.149	.062	.06	.097	.210	.168	.000		
	N	113	113	113	113	113	113	113	113	113	
10. Semantic fluency *Gr*	Pearson correlation	−.170	.175	.382**	.239**	.159	.087	.289**	.049	.172	—
	Sig. (two-tailed)	.074	.066	.000	.005	.094	.180	.001	.302	.070	
	N	113	113	113	113	113	113	113	113	113	113

Note. AUT = alternative uses task; AQT = alternative questions task; *Gf* = fluid intelligence; *Gr* = retrieval ability; Sig. = significance.
*$p < .05$. **$p < .01$.

Received January 30, 2023
Revision received September 20, 2023
Accepted September 23, 2023 ■