CHAPTER 2

The Trouble with "Creativity"

John Baer

As a creativity trainer I assumed creativity was domain general and that I was teaching techniques that would enable students to be more creative in whatever they did. When as a creativity researcher I tried to prove that, however, the evidence eventually forced me to accept the domain specificity of creativity and to understand what domain specificity implies for creativity training, as well as for creativity research, theory, and assessment. Because the content of creative products and the processes that produce those artifacts vary so much by domain, the term "creativity" has little meaning as a general concept. Understanding that what we lump together as "creativity" is a thin abstraction that tells us little about actual creativity will result in better creativity research, theory, training, and assessment.

More than anything else, my creativity research has focused on the question of whether the skills, approaches, dispositions, heuristics, talents, and any other factors that might underlie creativity are domain general or domain specific. That was not where my work in creativity started, however, nor where I thought it would lead. I've always been interested in creativity – Who isn't? – but my involvement in the *research* side of creativity grew, unexpectedly, out of the creativity training was doing decades ago using the CPS model.

CPS is an all-purpose approach to solving challenging problems that originated in the work of Alex Osborn (1953), the advertising executive who invented brainstorming. CPS can be used to solve tough problems of any kind. It was great fun, it seemed to work well (Baer, 1988), and I enjoyed teaching it.

Then I read Howard Gardner's (1983) seminal book on multiple intelligences theory, *Frames of Mind*. I was more interested in creativity than intelligence, but if what Gardner was saying about the modularity of intelligence was true, might it not also be true of creativity? A modular creativity would threaten the underpinnings of CPS, because if creativity varies from domain to domain – if the skills underlying creativity in the visual arts and

creativity in writing and creativity in math or science are different from and essentially independent of each other – then how could CPS training work across all domains? I decided to test this idea.

A research design to test the modularity of creativity seemed obvious. I would simply assess participants' creativity in different fields and then show (as I hypothesized) that (1) there was considerable overlap and that (2) creativity in one area predicted creativity in other areas. If the same people tended to be creative in many domains – if tasks X and Y came from different "intelligences" as defined by Gardner and there was a substantial correlation between creativity in X and creativity in Y – then it would show that creativity was not modular. As Ivcevic (2007) nicely summarized this idea years later:

> Domain generality would be supported by high intercorrelations among different creative behaviors . . . while domain specificity would be supported by relatively low correlations among different behaviors. (p. 272)

This test of domain generality/domain specificity – looking for intercorrelations across domains – is precisely the test that critics claim multiple intelligences theory, which argues for modularity in intelligence, has failed. In 1994 the Board of Scientific Affairs of the American Psychological Association (APA) established a Task Force to produce "an authoritative report" (Neisser et al., 1996, p. 77) on what is actually known about intelligence. The Task Force concluded that "subtests measuring different abilities tend to be positively correlated: people who score high on one such subtest are likely to be above average on others as well" and that psychometric evidence suggested "a hierarchy of factors with *g* at the apex" (p. 78).

This was exactly how I (and probably most people in the field of creativity theory; see, e.g., Amabile, 1983) thought of creativity. There were certainly domain-based differences, but there was also a large, domain-transcending core to creativity. My goal was to prove this, using the same approach that defenders of *g* used in arguing against multiple intelligences. Having demonstrated the domain generality of creativity (as I assumed my study would do), I might next assess the creativity, in different domains, of participants in CPS workshops before and after training. This, I trusted, would show that CPS boosted creativity across domains.

As I envisioned this program of research, however, I quickly hit a roadblock: How to assess creativity in different domains? I was interested in measuring actual creative performance. All of the Ps in the 4P Model (Rhodes, 1961, 1987) might be part of creativity, but three of these Ps (person, process, and press) are really there in the service of the fourth, product.

A press (environment) that supports creativity, a thinking process (or any other kind of process, such as the CPS model) that leads to creativity, and a personality that is conducive to creativity are creativity-relevant only if they tend to be associated with creative performances or products. If an environment, process, or personality was no more likely than chance to produce creative outcomes, then by definition it would not be associated with creativity. I therefore was interested in creative products or performances, even though my goal was related to a creative process (CPS).

Most of the available measures of creativity tended to focus on creative processes or personality traits, and they all had an even more significant drawback: They assumed domain generality. For example, the Torrance Tests, which were by far the most widely used creativity assessments available, assumed domain generality, despite the fact that they came in two versions, figural and verbal (Kaufman, Plucker, and Baer, 2008; Torrance and Presbury, 1984). As Plucker (1998) argued:

> No assumption is made that performance is specific only to the task or content area addressed in a particular divergent-thinking test. Even the creation of figural and verbal versions of the TTCT is not an acknowledgment of the possibility that creativity is content general. (p. 179)

All other major creativity tests had the same problem.

Luckily, 1983 also marked the publication of Amabile's game-changing *The Social Psychology of Creativity*, in which she introduced the consensual assessment technique (CAT). The CAT has been called the "gold standard" of creativity assessment (Carson, 2006) because its validity is rooted in actual creative performance and is determined in the same way that creativity is discerned in the "real world": via the consensus of experts in the domain. Nobel Prizes, Academy and Screen Actors Guild Awards, the Pulitzer and Booker Prizes, the Breakthrough Prizes, and many other major awards are decided by experts in the field in question. One might wonder how else could such honorees be chosen: by applying a rubric designed by psychologists? Hardly. How does one tell if a work of art or a scientific theory or a musical composition is a work of genius? Other than asking people who should know – the experts in the domain – there really aren't any good options. Experts may change their decisions (and their criteria) over time, of course, but at any moment in time, it is the consensus of experts in a field that determines what work is creative and what is pedestrian. The CAT works essentially the same way.

And there was one more feature of the CAT that made it perfect for the studies I wanted to do. Not only is the CAT probably the best available

measure of creativity; it is also (unlike every other major test of creativity available at the time) noncommittal about the question of domain generality v. domain specificity. CAT assessments are based on specific tasks, such as writing a poem or a story or making a collage (the three most widely used tasks, although many others have also been used). Whether the creativity ratings obtained with the CAT are simply valid assessments of creativity in the domain of the task (such as poetry or art) or of creativity more generally is something the CAT neither assumes nor predicts. These are open empirical questions that the CAT takes no stance on. Some researchers (e.g., Amabile herself; Amabile, 1983, 1996) have interpreted CAT scores without reference to domain, assuming that the specific tasks used would not influence a study's outcome, while others have used them only to assess creativity in specific domains (e.g., Baer, 1991, 1994, 1996).

The results of these studies were consistent, both in my research and the work of others. As Plucker (1998) noted, "the conclusions of researchers using the CAT are almost always that creativity is predominantly task or content specific" (p. 181). The across-domain correlations have been vanishingly small. These studies have provided no evidence of domain generality; on the contrary, they have consistently supported domain specificity.

My initial studies therefore led me to reverse my predictions for later studies as it became increasingly clear that the data simply didn't leave much room for theories of domain-general creativity. Unlike intelligence, in which intercorrelations among assessments in different domains have shown a significant overlap, creativity was showing itself to be quite domain specific. In fact, the size of the domains was even smaller (making the number of relevant domains considerably larger) than Gardner's eight intelligences. Creativity, research showed, was much more like expertise (which is very domain specific) than intelligence (in which there is substantial evidence for *g*). I've reviewed all of this work, both mine and others', in Baer (2016) and will only present the results from three representative papers here (Baer, 1991, 1994, 1996). In all three of these (as well as many others that preceded and followed these three) all artifacts were judged using the CAT. Details of the measures employed in each domain can be found in the original papers.

The 1991 paper discussed four studies using participants of different ages. Participants in one group of studies were eighth-grade students who created four artifacts: writing a poem, writing a story, creating an interesting mathematical equation, and writing an interesting math word problem. The within-domain tasks (poetry- and story-writing in the verbal domain; inventing interesting math equations and writing interesting math

Table 2.1 *Intercorrelations among creativity ratings (raw scores)*

Task	Poetry	Story	Word problem	Equation
Poetry	–	0.23	0.31*	−0.14
Story	–	–	0.20	−0.03
Word problem	–	–	–	−0.20

* $p < 0.05$

word problems in the math domain) were designed to be quite different from each other even though they nominally fell in the same domain. As Tables 2.1 and 2.2 show, cross-domain correlations were low, and even tasks within the same larger domain showed little evidence that they were rooted in the same sets of skills.

The other three studies reported in the 1991 paper had similar outcomes using participants of different ages: second-grade students, fourth-grade students, and adults. The cross-domain correlations were low in all groups, including a second testing of the fourth-grade students a year later as fifth graders. (This year-later testing was not done with the other groups.) The only statistically significant correlations between creativity ratings in that study were between scores obtained on the same task in fourth and fifth grades. The same students who were more creative on a given task in fourth grade tended to be more creative on that task in fifth grade, indicating a consistency of creative performance within a domain over time.

But what about creativity training, the question that got me started in creativity research? In the 1994 study I trained students in the experimental group using standard divergent thinking exercises (a major component of CPS training) using a variety of topics. All the divergent thinking activities were of a verbal nature. I trained the control group in ways of solving mathematical word problems with no divergent thinking activities.

Table 2.2 *Intercorrelations among creativity ratings; variance attributable to IQ removed*

Task	Poetry	Story	Word problem	Equation
Poetry	–	−0.01	0.19	−0.14
Story	–	–	0.05	0.07
Word problem	–	–	–	−0.45*

* $p < 0.05$

All subjects were then given five tasks: telling stories, writing stories, writing poems, writing mathematical word problems, and making collages. Experts evaluated the creativity of each product. The divergent thinking groups scored significantly higher than controls on the storytelling, story-writing, and poetry-writing tasks. The lack of correlations among scores on the five tasks, however, suggests that several task-specific factors, rather than one general factor, led to observed group differences. This is consistent with previous research using subjects untrained in divergent thinking in showing that divergent thinking is not a general trait (Baer, 1994, p. 35).

The divergent thinking training appeared to have an effect, but not a general effect:

> Although divergent thinking does not appear to be either a single skill or a distinct set of skills widely applicable within broad cognitive domains . . . what is commonly referred to as divergent thinking may describe a large constellation of skills, each influencing creative performance on different tasks. (p. 43)

The 1996 study was a different kind of training study. Rather than teaching divergent thinking skills in what might be called a shot-gun approach, I targeted the training by using only a very narrow range of content for the exercises. The participants were seventy-nine seventh-grade students in the experimental group and a matched group of control students in the same school. Random selection determined groups. The experimental group students were trained over several sessions using only poetry-relevant divergent thinking exercises. Both groups later wrote both poems and stories. These were rated for creativity using the CAT by experts who worked independently and who did not know who had had the training sessions and who had not been trained. The students with the poetry-relevant divergent thinking training wrote more creative poems than the control group, but their short stories were no more creative than those of untrained subjects.

Even divergent thinking, I discovered, is very domain specific. The training made a clear difference, but only in areas that matched the training. This is similar to efforts to increase cognitive skills through training. It can be done, but the effect is very narrow with little or no transfer. Wishful thinking that practicing one kind of cognitive skill will result in across-the-board cognitive improvements probably accounts for the recent popularity of brain-training programs like Luminosity, Jungle Memory, and CogniFit (Day, 2013), even though there is no evidence supporting such cross-domain transfer (Katsnelson, 2010; Owen et al., 2010; Redick et al., 2013; Thompson et al., 2013). Divergent thinking training seemed to follow

the same pattern: significant effects, but only on tasks similar to those used in training, with little or no impact even on different kinds of tasks within the same larger domain. The need for fairly extreme domain specificity to assess the effects of creativity training echoes what Pretz and McCollum (2014) wrote about the need for extremely domain-specific analyses: "Perhaps prior studies of domain-specific creativity were not specific enough" (p. 233).

Other researchers have conducted similar studies and have also found very limited cross-domain correlations (e.g., Han, 2003; Han and Marvin, 2002; Runco, 1987, 1989; Ruscio, Whitney, and Amabile, 1998). A convergence of research results by a diverse group of researchers is important, of course. It is especially so when it comes from researchers trying to prove you wrong, as happened in a large 1996 study by Conti, Coon, and Amabile that combined the results of three previously reported studies with overlapping participants (ninety young adults enrolled in an Introductory Psychology course).

Conti, Coon, and Amabile's (1996) participants completed a total of four story-writing tasks and three art activities. Conti, Coon, and Amabile predicted positive cross-domain correlations, in contrast to the prediction of domain-specificity theory. (Recall Ivcevic's [2007] formulation: "domain specificity would be supported by relatively low correlations among different behaviors"; p. 272.) Conti, Coon, and Amabile made these predictions, in fact, in direct response to my research:

> In contrast to recent proposals by Baer (1991, 1993, 1994), the componential model predicts that because there are cross-task skills that contribute to creativity, creativity measures will be positively correlated across different tasks and situations. (p. 386)

There were a total of thirteen cross-domain correlations in Conti, Coon, and Amabile's study, and there was simply no evidence of domain generality in these correlations at all. Of the thirteen correlations of this kind, *none* – not one of the thirteen – was statistically significant. Even mean scores from the seven tasks in the two domains – the mean of the four writing tasks and the mean of the three art tasks – did not produce a statistically significant result. (As both domain specificity and domain generality predict, the within-domain correlations were strong and statistically significant, but this outcome is irrelevant to the generality/specificity question.)

Amabile, the CAT's creator, has argued for both domain-specific and domain-general factors in creative performance (Amabile, 1983, 1996). Her research, however, actually supports domain specificity rather powerfully

(Conti, Coon, and Amabile, 1996). But that was clearly not her intention in creating the CAT. She has maintained that the CAT itself is neutral on the question of specificity/generality and can be used, in exactly the way I proposed, to test for domain generality and domain specificity (as she did in the study just reviewed). Domain generality theorists have acknowledged that divergent thinking tests like the Torrance Tests assume domain generality, but have suggested that the CAT, in turn, assumes domain specificity (e.g., Plucker, 1998). As Amabile herself has argued, however, that is simply untrue. In fact, in much of her research with the CAT, Amabile (1983, 1996) herself has treated CAT scores as general measures of creativity, not measures of creativity in a particular domain.

Depending on one's measurement goals and theoretical stance, one might think of the CAT as either a domain-general or a domain-specific test – or at least one could, until research showed, consistently, that the CAT could only measure creativity in the specific domain from which the artifacts being judged are drawn, because that is the only kind of creativity there is. To the extent that domain specificity is true, *all* tests of creativity are necessarily domain specific, regardless of their intent or their claims, because there is simply no domain-general factor to measure; all one can measure is creativity in the domain(s) included in the test. By including tasks from a variety of domains, one might create a test that measures creativity in several domains, but that would be, at best, a multiple-domain test, not a domain-general test of creativity.

Which brings me back to the Torrance Tests, which come in two versions, figural and verbal. The choice of which to use is simply a matter of convenience or suitability to the sample, because both are offered as domain-general tests (Plucker, 1998). These tests assume domain generality, but they have nonetheless provided evidence for domain specificity. They have done this in two ways: (1) by proving to be two independent and essentially uncorrelated measures, and (2) by evidencing mutually contradictory results in validation studies.

(1) According to one of his closest collaborators, Torrance himself found that the figural and verbal Torrance Tests were measuring two very different sets of skills.

> Responses to the verbal and figural forms of the TTCT are not only expressed in two different modalities . . . but they are also measures of different cognitive abilities. In fact, Torrance (1990) found very little correlation ($r = .06$) between performance on the verbal and figural tests. (Cramond, Matthews-Morgan, Bandalos, and Zuo, 2005, pp. 283–284)

It's rather hard to argue that two tests are measuring the same construct if their shared variance totals less than one-half of 1 percent, is it not? This is exactly the approach I used when I hoped to find domain generality in creative performance, and it is exactly the approach that the Board of Scientific Affairs of the APA's Task Force on Intelligence (Neisser et al., 1996) used as a primary way to show that intelligence has a significant domain-general component. Domain generality failed that test using CAT measures, as described above. And domain generality also failed that test using the two most widely employed divergent thinking tests. Plucker (1998) has argued that "[p]erformance assessments produce evidence of task specificity, and creativity checklists and other traditional assessments suggest that creativity is content general" (p. 180), but even these most traditional of all creativity assessments – the Torrance Tests of Creative Thinking – have provided strong evidence for domain specificity. By offering two versions of the Torrance Tests – even though they may have been intended to measure a single, domain-general set of abilities – the Torrance Tests have in fact allowed an unexpected test of domain generality. That test, conducted by Torrance himself, failed to find evidence of domain generality (with a correlation of just 0.06). As Sawyer (2012) concluded, "Different tests, each designed to measure creativity, often aren't correlated with one another, thus failing to demonstrate convergent validity" (p. 61).

(2) Validation studies of the Torrance Tests have received mixed reviews over the years. Many have concluded, as did Kogan (1983), Wallach, (1970), Anastasi (1982), and Crockenberg (1972) many years ago and as many reviewers have done more recently (Baer, 1993, 2011a, 2011b; Sawyer, 2012; Simonton, 2007; Sternberg, 1985), that no domain-general divergent thinking test has been validated as predictive of creativity. One common criticism of studies that have claimed to show validity of the Torrance Tests revolves around the kinds of criterion variables Torrance used as indicators of creativity, such as changing religious affiliation or subscribing to a professional journal. (On what basis might one conclude that things like changing religious affiliation or subscribing to a professional journal evidence creativity? There is no clear answer.) And then there is the problem that all of these indicators are based on self-report, which brings in another source of validity concerns. Crockenberg (1972) argued that "given the creativity criteria used . . . [the results of these studies] should not be taken too seriously" (p. 35) and Sternberg (1985) argued that "such tests capture, at best, only the most trivial aspects of creativity" (p. 618). More recently Sawyer (2012) pointed out that Guilford himself admitted that divergent thinking tests don't correlate highly with real-world creative output, adding

that "although there remain some dissenters, most psychologists now agree that DT tests don't predict creative ability" (p. 51).

Despite these problems, validation studies of the Torrance Tests offer some unexpected (and unsought) evidence of domain specificity of divergent thinking. Plucker (1999) chose to reanalyze data from the Torrance validation study that provided the "most compelling" evidence for validity of the Torrance Tests, arguing that "[a]ny analysis of this topic should begin with this seminal study" (p. 104). Two hundred students took the Torrance Tests every year in grades 1–6. Torrance then used these divergent thinking test scores to predict a variety of self-reported criterion measures. As noted above, there has been much dispute about the validity of the criterion measures Torrance used, but Plucker's paper was interesting because he determined that in his reanalysis of Torrance's longitudinal data, one of the two Torrance Tests that subjects had taken did positively predict later self-reported creative performance – but the other did not. The difference in the predictive ability of the two Torrance Tests is telling, and Plucker couldn't explain why one of the Torrance Tests – the Verbal Test – predicted creativity while the other Torrance Test – the Figural Test – did not without resorting to domain specificity:

> The importance of verbal DT relative to figural DT may be due to a linguistic bias in the adult creative achievement checklists. For example, if a majority of the creative achievements required a high degree of linguistic talent, as opposed to spatial talent or problem solving talents, the verbal DT tests would be expected to have a significantly higher correlation to these types of achievement than other forms of DT. (Plucker, 1999, p. 110)

Exactly. Domain-specificity theory predicts that different measures of creativity rooted in different domains will predict creative performance *only in their respective domains*.

Domain generality theorists like Lubart and Guignard (2004) (who argued that "performance-based evaluations provide results favoring a domain-specific view"; p. 53) and Plucker (1998) agree that performance assessments routinely support domain specificity. They claim that creativity checklists provide evidence of domain generality, however. Do they?

Plucker (1998) cited a study by Runco (1987) that used students' self-reported levels of creativity in seven performance domains:

> Runco (1987) compared students' creativity checklist responses to quality ratings of the students' creativity (scored using a technique not unlike the CAT). The students' checklist scores provided evidence of content generality, and the quality ratings suggested content specificity. (p. 181)

So self-reported creativity by these students suggested domain generality, but their actual creative performance argued for domain specificity. Which kind of measure should we trust? (This is reminiscent of the Marx brothers' line from *Duck Soup*, "who are you going to trust, me or your lying eyes?")

There have been two excellent reviews of self-report creativity checklists in recent years, and it seems that the best one can say about such checklists is that they *might* have *limited* validity, especially some of the newest checklists, *but only when used to make very low-stakes decisions*. When Reiter-Palmon, Robinson, Kaufman, and Santo (2012) reviewed several frequently used self-report measures of creativity, they urged cautioned:

> These results suggest that although self-perceptions of creativity may provide some information about creativity, researchers should be cautious when using this measure as a criterion. (p. 107)

But when Silvia, Wigert, Reiter-Palmon, and Kaufman (2012) reviewed "four new and promising [creativity self-report] scales" (p. 19), they concluded that, although most such reviews "end on a grim note" (p. 31), the four new scales were more promising, at least if used only for low-stakes assessment.

These are hardly ringing endorsements, and because most of the creativity-checklist research that has been cited to support domain generality did not use the four "new and promising" scales, what self-reported creativity checklists can tell us about domain generality or domain specificity is probably quite limited.

As Sawyer (2012) concluded in his textbook *Explaining Creativity*, "[a] wide range of studies has shown that much of creative ability is domain-specific" (p. 60). Is there any domain generality at all? Probably. In our APT model of creativity, Kaufman and I proposed a hierarchy with some, possibly minor, domain-general factors (such as intelligence) and several levels of increasingly domain-specific factors that do most of the creativity heavy lifting (Baer and Kaufman, 2005; Kaufman and Baer, 2004).

Does this mean that CPS, which is where I started my creativity research journey, should be abandoned? Not at all. But it means that we must recognize that CPS (and divergent thinking skills, which are integral to CPS) will work differently (and require different training and practice) in different domains. Consider this parallel: We recognize that acquiring expertise generally requires practice and/or study, but we don't assume that the kinds of practice or study that lead to expertise in music will lead to expertise in cosmology. Expertise is very domain specific. So is creativity.

Confused (and often unrecognized) assumptions of domain generality, which are unfortunately common, have made it more difficult to understand creativity and have led to a plethora of conflicting and impossible-to-replicate results. For example, consider the relationship (or lack of relationship) between creativity and mental illness. Much ink has been spilled in this dispute going back many decades, but we now know that the presence or absence of such a linkage, as well as its degree where it exists, depends on the domain in question.

> The rate and intensity of adulthood symptoms vary according to the particular domains in which creative genius is expressed . . . geniuses in the natural sciences tend to be more mentally healthy than in the social sciences; geniuses in the social sciences, more so than those in the humanities; and geniuses in the humanities, more so than those in the arts. (Simonton, 2010, pp. 226–228)

Assuming domain generality made the truth about the creativity-mental illness connection impossible to see. Only a domain-specific orientation made it possible to undercover the truth.

Does this make studying and training creativity less exciting? Probably. Believing that one's students will become more creative in everything they do is a more satisfying theory for teaching CPS. But it makes creativity training even more essential (and more time consuming) when one realizes that it must be done not once and for all, but domain by domain. Ditto for creativity research and theory. Believing that one's theory is about all of creativity is more motivating than understanding that one's theory will probably only work in some domains (and will need to be tested in each). But assuming domain generality has a serious danger: It means one's research probably won't withstand replication when tried in other domains.

In some ways the root of the domain-specificity problem goes back to a dispute in philosophy at least as old as Plato and Aristotle. Are forms – abstract ideas like "beauty," "quality," and "goodness" – real, with an existence outside our minds? Would these abstractions be part of the world whether or not humans ever noticed them? Or are such concepts simply words that have no objective correlate, no independent existence, in the real world? Would there exist such a thing as "beauty" in the world even if no one ever noticed it, or is "beauty" a human invention? Nominalists argued that we must "avoid the temptation . . . of assuming that the ontological structure of the world matches the structure of our thoughts" (Kronman, 2016, p. 347).

There are times when it is helpful to use the term "creativity" to describe, collectively, a set of things from diverse domains. (I won't venture a definition beyond noting that novelty and appropriateness, as defined by the domain in question, seem to be key ingredients.) But the usefulness of the concept "creativity" does not mean that there actually exists in the world something that corresponds to what we intend when we use this term. The abstractions "beauty," "quality," and "goodness" can also be useful terms, but most of us can use them without losing sight of the fact that these are actually collections of things that are, in their actual manifestations, quite varied and for the most part unrelated. Beauty, quality, and goodness are all things that matter to us, but they are not things that are readily amenable either to research or to training – *except* research and training that focuses on beauty, quality, or goodness *in specific domains*.

We need a more nominalist understanding of "creativity," one that recognizes that the actual ideas and things we describe as creative – and the processes that yield those creative ideas and artifacts – are for the most part unrelated unless those things and ideas come from the same domain.

It is those domain-specific manifestations and domain-specific processes that we should be studying and training.

REFERENCES

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997–1013.
(1983). *The social psychology of creativity*. New York: Springer.
(1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.
Anastasi, A. (1982). *Psychological testing*. New York: Macmillan.
Baer, J. M. (1988). Long-term effects of creativity training with middle-school students. *Journal of Early Adolescence*, 8, 183–193.
(1991). Generality of creativity across performance domains. *Creativity Research Journal*, 4, 23–39.
(1993). *Creativity and divergent thinking: A task-specific approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
(1994). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal*, 7, 35–46.
(1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior*, 30, 183–187.
(2011a). Four (more) arguments against the Torrance Tests. *Psychology of Aesthetics, Creativity, and the Arts*, 5, 316–317.
(2011b). How divergent thinking tests mislead us: Are the Torrance Tests still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts*, 5, 309–313.

(2016). *Domain specificity of creativity*. San Diego, CA: Academic Press/Elsevier.
Baer, J., and Kaufman, J. C. (2005). Bridging generality and specificity: The Amusement Park Theoretical (APT) model of creativity. *Roeper Review*, 27, 158–163.
Carson, S. H. (2006). Creativity and mental illness. Invitational Panel Discussion Hosted by Yale's Mind Matters Consortium, New Haven, CT, April 19.
Conti, R., Coon, H., and Amabile, T. M. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies. *Creativity Research Journal*, 9, 385–389.
Cramond, B., Matthews-Morgan, J., Bandalos, D., and Zuo, L. (2005). A report on the 40-year follow-up of the Torrance Tests of creative thinking. *Gifted Child Quarterly*, 49, 283–291.
Crockenberg, S. B. (1972). Creativity tests: A boon or boondoggle for education? *Review of Educational Research*, 42, 27–45.
Day, E. (2013). Online brain-training: Does it really work? *The Guardian*, April 20. www.theguardian.com/science/2013/apr/21/brain-training-online-neuroscience-elizabeth-day.
Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
Han, K. S. (2003). Domain specificity of creativity in young children: How quantitative and qualitative data support it. *Journal of Creative Behavior*, 37, 117–142.
Han, K. S., and Marvin, C. (2002). Multiple creativeness? Investigating domain-specificity of creativity in young children. *Gifted Child Quarterly*, 46, 98–109.
Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *Journal of Creative Behavior*, 41, 271–290.
Katsnelson, A. (2010). No gain from brain training. *Nature News*, 464(7292), 1111.
Kaufman, J. C., and Baer, J. (2004). The Amusement Park Theoretical (APT) model of creativity. *Korean Journal of Thinking and Problem Solving*, 14(2), 15–25.
Kaufman, J. C., Plucker, J. A., and Baer, J. (2008). *Essentials of creativity assessment*. New York: John Wiley.
Kogan, N. (1983). Stylistic variation in childhood and adolescence: Creativity, metaphor, and cognitive styles. In *Handbook of child psychology: Vol. 3. Cognitive development*, 4th ed., edited by P. H. Mussen, 628–706. New York: John Wiley.
Kronman, A. T. (2016). *Confessions of a born-again pagan*. New Haven, CT: Yale University Press.
Lubart, T., and Guignard, J.-H. (2004). The generality-specificity of creativity: A multi-variant approach. In R. J. Sternberg, E. L. Grigorenko, and J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 43–56). Washington, DC: American Psychological Association.
Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., and Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77–101.
Osborn, A. F. (1953). *Applied imagination*. New York: Charles Scribner's Sons.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., Howard, R. G., and Ballard, C. G. (2010). Putting brain training to the test. *Nature*, 465(7299), 775–778.

Plucker, J. A. (1998). Beware of simple conclusions: The case for the content generality of creativity. *Creativity Research Journal*, 11, 179–182.

(1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal*, 12, 103–114.

Pretz, J. E., and McCollum, V. A. (2014). Self-perceptions of creativity do not always reflect actual creative performance. *Psychology of Aesthetics, Creativity, and the Arts*, 8, 227.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., Kane, M. J., and Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, 142(2), 359.

Reiter-Palmon, R., Robinson, E., Kaufman, J., and Santo, J. (2012). Evaluation of self-perceptions of creativity: Is it a useful criterion? *Creativity Research Journal*, 24, 107–114.

Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42(7), 305–310.

(1987). An analysis of creativity. In *Frontiers of creativity research: Beyond the basics*, 216–222. Buffalo, NY: Bearly.

Runco, M. A. (1987). The generality of creative performance in gifted and nongifted children. *Gifted Child Quarterly*, 331, 121–125.

(1989). The creativity of children's art. *Child Study Journal*, 19, 177–190.

Ruscio, J., Whitney, D. M., and Amabile, T. M. (1998). Looking inside the fishbowl of creativity: Verbal and behavioral predictors of creative performance. *Creativity Research Journal*, 11, 243–263.

Sawyer, K. (2012). *Explaining creativity: The science of human innovation*. 2nd ed. Oxford: Oxford University Press.

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

Simonton, D. K. (2007). Creative life cycles in literature: Poets versus novelists or conceptualists versus experimentalists? *Psychology of Aesthetics, Creativity, and the Arts*, 1, 133–139.

(2010). So you want to become a creative genius? You must be crazy! In *The dark side of creativity*, edited by D. Cropley, J. Kaufmann, A. Cropley, and M. Runco, 218–234. New York: Cambridge University Press.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49, 607–627.

Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., Chang, P., Pollard, K., Lala, N., Alvarez, G. A., and Gabrieli, J. D. E. (2013). Failure of working memory training to enhance cognition or intelligence. *PloS One*, 8(5), e63614. www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0063614.

Torrance, E. P. (1974). *Torrance Tests of creative thinking: Norms-technical manual*. Lexington, MA: Ginn.

(1990). *The Torrance Tests of creative thinking: Norms-technical manual*. Bensenville, IL: Scholastic Testing Service.

Torrance, E. P., and Presbury, J. (1984). The criteria of success used in 242 recent experimental studies of creativity. *Creative Child and Adult Quarterly*, 9, 238–243.

Wallach, M. A. (1970). Creativity. In *Carmichael's handbook of child psychology*, vol. 1, 3rd ed., edited by P. H. Mussen, 1211–1272. New York: John Wiley.