Introduction: Creativity Matters, so How We Think about It Matters

Is There a Creativity Crisis?

In July 2010 Newsweek magazine’s cover announced that America was suffering a “Creativity Crisis.” This was illustrated with a line-up of children’s crayons arrayed like a plummeting graph. An ominous subheading spelled out the danger: “For the first time, research shows that American creativity is declining.”

This is nonsense. As far as research can possibly determine such things—which is, I’m sad to report, not very far at all—there is not, and there was not in 2010, a crisis of any kind in American creativity.

The sensationalistic Newsweek cover was based on a reported twenty-year decline in scores on one very old test of creativity (Kim, 2011). Although once widely used, this is a test whose validity has been challenged for almost half a century.

Other measures of creativity, it turns out, show an increase in creativity over the same period of time that the supposed decline was happening (Gardner & Davis, 2013; Weinstein et al., 2014). But those results, which point in the opposite direction, are equally meaningless in regard to overall American creativity.

There is simply no way to know if creativity is waxing or waning in the USA as a whole. The question, as framed, has no possible answer, because there is no way to sum creativity in sculpture and creativity in cosmology, creativity in cooking and creativity in sports, creativity in interpersonal relations and creativity in poetry.

This contradiction—one set of creativity tests showing a decline, the other an upswing over the same period—is not surprising to anyone who knows much about what pass for “creativity tests.” These tests routinely disagree with each other, even tests made and sold by the same test developers (including the one that showed the twenty-year decline). None of these tests can really tell us much about creativity in general.

So, in 2010 there was no reason, based on the research reported in the Newsweek story, to sell off one’s portfolio of US stocks. (Disclaimer: Expertise—like creativity, as I will explain—is very domain specific. Being an expert in one domain does not make one an expert in other, unrelated domains. I’ve studied creativity for decades, have published scores of books and research papers about creativity, and have even won some awards for that work, so I can claim some expertise in this area. But I’m not an economist and have no expertise as a financial advisor. I know that there is no reason to believe...
there has been a general decline in American creativity, but I have no idea if there may have been other reasons to sell, or to buy, US securities back in 2010. Or at any other time.)

There is no creativity crisis, and yet there is a very real and important crisis in creativity. This crisis is not about tumbling scores on one (of many available) so-called “creativity tests,” such as the one that the Newsweek magazine focused on. Nor is it about a sweeping, or even a slight, overall change in actual creative performance by either children or adults. There are changes over time, of course, but they are much more limited in scope. In some areas there is evidence for a modest increase, in others an similarly modest decrease (see, e.g., Gardner & Davis, 2013, and Weinstein et al., 2014, for examples of both increases and decreases in creativity over time in specific areas, and see Barbot & Said-Metwaly, 2021, for a more thorough explanation of where the research that was the basis for the Newsweek article went awry).

The crisis is in how we (mis)understand creativity – a misunderstanding that leads to poor creativity training and unreliable creativity assessment, both based on weak creativity theory and confused (and often contradictory) creativity research. It is the field of creativity studies that is in crisis, not actual creativity.

You may be breathing a sigh of relief and thinking, “So what? I care about actual creativity, not creativity research.” And, in a way, you’d be right to think this. It’s actual creativity we should care about. As long as their work doesn’t impinge on actual creativity, what harm is there if creativity researchers and theorists (like me) are confused?

But creativity – whether in the form of monumental advances in science, or a toddler’s epiphany; transcendent poetry, or a more tasty soufflé; solutions to world problems, or just a different way to complete the most mundane task – matters very much. Creativity matters for reasons ranging from the need to solve international problems that threaten the health of the planet to the simple joy that creativity, even creativity in small, pedestrian doses, brings to our lives. If we want to nurture creativity in children and adults – and a constant stream of books, TED talks, and opinion pages suggests that we do – we need better creativity theory and research.

In an interesting twist, Newsweek’s competitor Time also had a cover story about creativity, this one eleven years later. The schoolchildren whose alleged twenty-year decline in creativity that Newsweek reported in 2010 would thus be young adults at the time of Time’s June 28, 2021, cover story. These twenty- and thirty-year-old creative dullards, graduates of the “creativity crisis,” were at just the right age in 2021 to bring that predicted crisis to a head in the real world. But the Time cover announced an “Innovation Boom,” which is pretty much the opposite of a creativity crisis. “We’re entering a new era of innovation,” the
story proclaimed. As it happened, the *Time* magazine story was about innovation in a specific domain – public health – not creativity in general. This, as we’ll see, is exactly how we should be thinking about creativity, domain by domain. But that’s getting ahead of our story.

**Unusual Uses for a Brick**

Unfortunately, most creativity theories (and the creativity research and creativity training based on those theories) are still stuck in Eisenhower-era models. Creativity research keeps de-bunking, and then modifying but essentially reinventing, the same theories, the same models, the same training, the same tests.

That creativity test that showed a twenty-year decline in scores? It was born in the 1950s. A half century ago, when the test was still relatively new, critics like University of California professor Susan Crockenberg were already arguing that the validity studies used to support that particular test (the Torrance Test of Creative Thinking) “should not be taken too seriously” (Crockenberg, 1972, p. 35).

A decade later a broader complaint emerged about the entire class of tests of which the *Newsweek* magazine test was just one example. These tests were (and are) all actually tests of divergent thinking, or DT. Some theorists have argued that DT may be part of creativity, but as Runco and Acar (2012) correctly note, DT test scores are, at best, possible “indicators of creative talent” (p. 66), not measures of creativity. DT tests are often mistakenly regarded as tests of creativity, however. Runco and Acar emphasize that this is “not a tenable view” (p. 67), and they are right, but the most widely used DT tests nonetheless call themselves tests of creativity, so it is not hard to understand how this confusion might have arisen. In the discussion that follows I will not test the reader’s patience by continually making this distinction and will often refer to these tests as creativity tests, even though they are actually DT tests.

To understand the problem with DT tests as possible indicators or predictors of creativity one needs to understand what DT is. Fortunately, it’s not a complicated idea. DT is simply the production of many different and unusual responses to an open-ended question. To qualify, the ideas needn’t be good ideas. They just need to be original, and plentiful.

Coming up with lots of strange ideas, even totally unworkable, goofball ideas, surely *seems* like the kind of thing that should be related to creativity, doesn’t it? The theory is indeed a compelling one.

About the same time DT was first proposed by Joy Guilford as a kind of thinking associated with creativity, and just before the first tests to measure it were created (whose direct descendant was the one cited in the *Newsweek*
magazine “Creativity Crisis” story), Alex Osborn independently invented brainstorming as a way to produce a lot of it (DT, that is, even though that word wasn’t yet available to describe it).

My students sometimes conflate DT and brainstorming, but brainstorming is just one tool for producing DT. It’s like a shovel and a hole. We use a shovel to get a hole, but the two are very different things. Shovels are probably the most commonly used tool for making holes, and brainstorming the most commonly used tool for producing DT. But there are other ways to get both holes and DT. Just as we should not confuse shovels with holes, we need to distinguish the tool of brainstorming from the goal of DT.

Divergent-thinking tests pose tasks like making a list of interesting things one might do with a brick. Surely this requires creative thinking; as I said just a few paragraphs ago, the theory is a compelling one. The complaint? Divergent-thinking tests are fun, but they predict actual creativity – creativity in the arts, in the sciences, in problem-solving of all kinds – either poorly or not at all. They sometimes predict one kind of creativity but not others. DT of some kind may indeed be a component of some kinds of creative thinking, but the tests that have been designed to measure it simply don’t tell us much about creativity in general.

How could that be? Here’s an analogy: Spelling is a component of writing, but we wouldn’t say that spelling tests are writing tests. There might be a positive correlation – better spellers might be better writers, although I know of no study that shows that, so that’s just a guess – but surely such a spelling test wouldn’t be a very good measure of writing skill in general. And what if the spelling test only included words related to a single topic, such as sports? A test that measured how well one could spell sports-related words would be even less valid as a test of writing. It might predict to some degree one’s skill in writing about sports, but say nothing about one’s ability to write about, say, biology or cooking.

These may seem like extreme cases, but in fact most DT tests don’t measure DT in general, only DT in a particular domain. Even if they predict creativity (to a limited degree) in one domain – and it’s not clear that they do, because researchers have too often been asking the wrong questions about DT tests, so we don’t really know in which domains this might be true and in which it might not be – one thing is increasingly clear: They fail to predict creativity in unrelated domains.

American Psychological Association (APA) past-president Robert Sternberg summed up the emerging consensus four decades ago: DT tests, he argued, “capture, at best, only the most trivial aspects of creativity” (Sternberg, 1985, p. 618).
That opinion is widespread in the creativity research community today. As Sawyer put it in his popular creativity textbook, “most psychologists now agree that DT tests don’t predict creative ability” (Sawyer, 2012, p. 51).

The continuing use of tests that most psychologists don’t think are good predictors of creativity is why the APA sponsored a debate a few years ago on the question “Are the Torrance Tests Still Relevant in the 21st Century?” (The two Torrance Tests of Creative Thinking are the most widely used DT tests. One of them was the basis for the Newsweek magazine cover story.)

Note that the question that the APA asked was not along the lines of “How good are the tests?” or “When might they be useful?” but more on the order of “Are they any good at all?” and “Should they ever be used? (Full disclosure: I was the psychologist invited by the APA to argue for their nonrelevance. Probably not fair, then, for me to judge who “won” the debate. And as my counterpart in that debate is both a friend and a respected colleague, I wouldn’t answer anyway. But I think it is fair to say that in being asked to argue that the Torrance Tests are irrelevant and should therefore be both mistrusted and avoided, I had the much easier assignment.)

So why are such tests still used? Because nothing better has come along, at least nothing better that is both cheap and easy to use. And even though we’re not very good at measuring creativity – even though our efforts are mostly wasted, and even though we know that had we chosen a different creativity test we would likely have gotten totally different results (which should be a clear indication that our tests are invalid) – people nonetheless want to measure creativity. Almost desperately. After all, if something can’t be measured, how can researchers study it? And what does it mean to describe something as increasing or decreasing, or even as present or absent, if it can’t be measured?

With some trepidation, let me make a comparison to a very different kind of measure, IQ tests. IQ tests are very controversial (hence my apprehension in even mentioning them), and the validity of IQ tests is not a subject I want to broach. Except to say this: If the most widely used IQ tests were as varied in their results – if they disagreed with one another to the same extent – as the most widely used creativity tests, there would be no controversy about IQ test scores. None at all. Why not? Because no one would be using those IQ tests. And no one would think they were worth defending.

But creativity tests, as lame and as widely criticized as they are (and have long been), are still used, both in research and in identification for school programs. Why? Because creativity matters to us – as it should. So we want to find a way, we feel we must find a way, to measure it. We almost desperately want measures of creativity, because how else can we promote and research and even talk about creativity (or sell magazines proclaiming its decline)? But
desperation is not the kind of motivation that typically leads to good decision-making.

I understand both the desire to measure creativity and the assumption that it must be measurable. Clearly some things are more creative than others, so it seems there must be some way to quantify those differences. But there is a real cost, and a real danger, in using such poor tests of creativity as the ones we have.

In computer science there’s a saying, “Garbage in, garbage out.” Nonsensical input data – or even just seriously flawed data that has at best marginal validity – produce nonsensical output. A different set of nonsensical or seriously flawed data would produce a different (equally nonsensical) result.

Sadly, one can prove almost anything one wants to prove about creativity simply by choosing the right test (and then someone else can disprove the same thing by using a different test).

I don’t mean to suggest that any creativity researchers – a group of people I respect greatly and am proud to count myself part of – are intentionally choosing tests that they know are either nonsensical or biased in order to falsely “prove” anything. But by unintentionally using tests of negligible validity – because that’s pretty all that has been available – creativity researchers often, totally inadvertently, produce research findings that are very different from those one might find had a different test been used (as in the case of the “Creativity Crisis,” where there are different measures of creativity that yield directly opposite results).

The problem is not with one test. Or with one theory, or one research program. If it were, psychological science would have replaced the bad tests, mistaken theories, and unproductive research programs with better ones. The problem is that creativity testing, creativity theory, and creativity research can’t seem to find anything better. All are stuck, often recycling the same tired ideas year after year, decade after decade. It may not be quite so dire as Dietrich and Kanso (2010, p. 845) have argued: “Ideas first proposed in the 1960s and 1970s – laterality, divergent thinking, low arousal, remote associations, or defocused attention – are still those, and in essentially the same form, that drive current research efforts.” It isn’t really quite that bad. But almost.

Philosopher of science Thomas Kuhn argued that new paradigms only get widely accepted after those who held earlier views leave the field. Physicist Max Planck put it more bluntly: “A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die” (quoted in Kuhn, 1970, p. 151).

But the theorists, researchers, and test makers who developed what might be called the standard model of creativity in the 1950s and 1960s have left the stage. At least two generations of researchers and theorists later, the field
remains stuck in the same ruts. It can’t be the intransigence of long-gone leaders of the field that is causing the lack of progress.

Sometimes what a scientific field needs to replace is not its leaders. It needs to replace its assumptions. Here’s an example: For centuries, scientists believed there was a luminiferous ether that allowed light waves to travel through empty space. Light was a wave, everyone agreed, and waves can’t pass through a vacuum, so space could not be a vacuum. But it is a vacuum, and understanding light required letting go of a long-held belief in the ether. Similarly, the assumption that planetary orbits must be circular led to some very creative, and complicated, models of our solar system. Progress, however, required ditching that (seemingly obvious) assumption and working with noncircular orbits.

Many smart and thoughtful people work in the field of creativity studies. Dedicated, curious, imaginative people who conduct ingenious studies designed to tease out creativity’s mystery. It’s not that we are learning nothing about creativity from their research. But they are hampered by an assumption about creativity that is simply false: the belief that creativity exists.

I hope you’re thinking, “Of course creativity exists!” There are creative theories in science (and creative ways to test those theories). There are creative works of art of all kinds. There are creative recipes, creative songs, creative ways to resolve disputes with friends, creative ways to play with children, creative ways to coach sports teams, and even creative ways to conduct psychological research about creativity.

The list of things that can be done creatively is, happily, endless.

When I say that creativity doesn’t exist, I should perhaps put scare quotes around it. The “creativity” that doesn’t exist is the “creativity” that researchers and theorists – and the rest of us as well – typically assume is behind all the creative things and ideas we encounter. A creativity that exists apart from (and is at the same time a part of) actual creative things, ideas, and performances.

“Creativity” is typically taken to mean something that runs through and unifies all those creative things, some special feature (newness? originality? beauty?) that all creative things share. Alternatively, “creativity” is sometimes thought of as a special kind of thinking process, a way of approaching problems, habit of mind, motivation, or personality trait that produces all those creative things and ideas, something shared by creative people (and the rest of us, but to a lesser extent) that makes it possible to go beyond the obvious, the mundane, the routine. Either way – be it a unifying essence or a unifying process that produces that essence – creativity is thought of as something separate and distinct from actual creative things and ideas.

But that kind of “creativity” doesn’t exist, even though there are, indeed, many creative things being produced every day, some world-altering, some
hardly noticed. (There are also many creative thoughts, some of which may never even get expressed but are creative nonetheless.) What is false is the idea – the generally accepted idea, both inside and outside the creativity research community – that there is something in the world, something other than actual creative things or ideas, that corresponds to the word “creativity.” And this false idea gets us into trouble.

Why This Matters

I used to be a creativity trainer. My classes lasted up to a week, and what I did – what I claimed I would do – was teach my students how to be more creative in anything and everything they did. I enjoyed those workshops, as did my students. I loved that I was making the world a better place by increasing the sum total of creativity and by making my students’ lives better by increasing their individual creativity, which I believed (and continue to believe) is one of the things that makes life worth living.

But then, in 1983, Howard Gardner published *Frames of Mind*, in which he argued that there was no such thing as general intelligence. Domain-specific intelligences, yes – things like logical/mathematical intelligence, verbal intelligence, interpersonal intelligence, spatial intelligence – but each operated only in its domain. No general intelligence.

Psychologists for the most part disagreed. Of course there were specific abilities, whether in the domains Gardner had identified or in other domains. But there were also some general abilities that were applicable and useful pretty much across the board. Their chief method of demonstrating this general, domain-transcending ability or set of abilities was simple: They showed that there were substantial positive correlations in measures of performance across different domains.

As I wrote earlier, disputes about intelligence, or intelligence testing, are not my concern here. Whether you agree with Gardner about the nature of intelligence or agree with his critics makes little difference in terms of the reality of “creativity.” But Gardner also suggested that creativity is likely to be domain specific in the same way, and although that was not the focus of his book, it was the part that I focused on because it challenged the possibility that the kind of creativity training I was doing was even possible. If all creativity was domain specific – if the thinking skills or ways to approach problems that led to creativity were different for every domain – then I had been selling my students a bill of goods. The kind of creativity I was teaching didn’t exist.

There were many responses to Gardner’s thesis, but the most compelling was the fact that if one measures abilities in a wide variety of domains, the
people who exhibit more intelligence in one domain tend to do so for most domains. Put another way, there were large positive correlations on measures of ability – the kinds of abilities generally thought of as showing intelligence – across domains.

So what I needed to do was show that creativity worked that way too. That creativity in any one domain predicted creativity in other domains. With a grant from the National Science Foundation, I set out to show this. At about the same time that Gardner’s book came out, Teresa Amabile had developed a very powerful way to measure creativity that focused on creativity in specific domains. Unlike DT tests and other available measures of creativity, which generally assumed that creativity is domain general, her method – the Consensual Assessment Technique – didn’t address the issue of domain specificity/generality at all, even implicitly.

The Consensual Assessment Technique was simply a way to assess creativity in a specific domain, and it could be used, at least theoretically, in any domain. How one interpreted the results of those tests was open. One could take them at face value as measures of creativity in the specific task and domain used in the assessment and therefore valid only as measures of creativity in that one domain. For example, if the task involved writing poetry, then the result could be interpreted as an indicator of poetry-writing creativity. If the task did not involve poetry but was, let’s say, an art-related task, then it could be interpreted as an indicator of artistic creativity and say nothing, one way or the other, about poetry-writing creativity.

Alternatively, one could assume that the task used and the domain of that task didn’t matter and that the creativity ratings based on any of these domain-based performances reflected creativity across the board, even though the task actually used in any of the assessments would necessarily come only from a single and very specific domain. Under such a domain-general interpretation, both a poetry-writing task and an art-creating task would actually be measuring the same thing: general creativity.

Amabile was interested in the effects of different kinds of motivation on creative performance. Addressing the specificity/generality question wasn’t at all her goal in developing this approach to creativity assessment (the details of which we will consider below). She assumed that whatever one learned about, say, the effect of intrinsic versus extrinsic motivation on a person’s creativity in collage-making would also be true of the effect of intrinsic versus extrinsic motivation on their creativity in other areas, like writing poetry. But because of the way her tests were designed, what she was actually measuring was creativity on a specific task, such as making collages, writing poetry, or writing stories (which were the three main tasks she used in her research).
Although not at all the purpose for which these tests were designed, they happened to be ideal for research on the domain-specificity/generality question. The appearance of this new method of creativity assessment just as Gardner’s unsettling book was published was perfect timing. (See Amabile, 1996, for a summary of the development of this approach to measuring creativity and the research she conducted with it. Unfortunately, the Consensual Assessment Technique is both very resource intensive and difficult to standardize, which often makes it difficult to use in many kinds of creativity research and in educational settings. See Baer, 2016, for a discussion of these issues.)

With Gardner’s challenge in mind and Amabile’s Consensual Assessment Technique in hand, I set out to show that creativity was highly domain general. Put another way, I expected to find that people who were highly creative when doing X were also more likely than chance to be creative when doing Y or Z. What I found, however, in study after study, was exactly the opposite of what I had hoped to find. Creativity wasn’t one thing; it was many, mostly unrelated things. More details about that initial research and the many follow-up studies that I and others conducted will come later. For now, let me be clear where we’re going: I’m arguing that there is no such thing as “creativity” that exists apart from actual creative ideas, creative products, or creative performances. I’m claiming that those creative ideas, creative products, and creative performances are not produced by a similar set of thinking processes, problem-solving approaches, personality traits, or general skills, and that they do not share any essence or essential feature that makes them creative.

**Philosophy: It All Goes Back to Plato**

The claim that there is no such thing as creativity is related to an important idea in philosophy. Plato (everything goes back to Plato, doesn’t it?) believed that there were individual things, like the mountain you plan to climb tomorrow, the courage with which you undertook some great challenge yesterday, and the chair you are sitting on today. No problem there. But Plato also believed there were things – ideal but very real things – that are the models for all actual, real-world mountains, instances of courage, and chairs, and it is these perfect models that give the less-than-perfect versions in which we encounter their essence.

Each mountain, each instance of courage, and each chair is independent and real, with particular features that distinguish it from other mountains, instances of courage, or chairs. But each (according to Plato) is based on a more universal, eternal, and unchanging *form*: an ideal and perfect mountain, an ideal conception of courage, an ideal chair. Everything in the world – every actual thing – is an imperfect version of a perfect form.