Large-Scale Assessments, Personalized Learning, and Creativity: Paradoxes and Possibilities

Ronald A. Beghetto
University of Connecticut

Abstract

Purpose: This article, based on an invited talk, aims to explore the relationship among large-scale assessments, creativity and personalized learning.

Design/Approach/Methods: Starting with the working definition of large-scale assessments, creativity, and personalized learning, this article identified the paradox of combining these three components together. As a consequence, a logic mode of large-scale assessment and creativity expressions is illustrated, along with an exploration of new possibilities.

Findings: Smarter design of large-scale assessments is needed. Firstly, we need to assess creative learning at the individual level, so complex tasks with high uncertainty should be presented to students. Secondly, additional process and experiential data while students are working on problems need to be captured. Thirdly, the human-artificial intelligence (AI) augmented scoring should be explored, developed, and refined.

Originality/Value: This article addresses the drawbacks of current large-scale assessments and explores possibilities for combining assessment with creativity and personalized learning. A logic model illustrating variations necessary for creative learning and considerations and cautions for designing large-scale assessments are also provided.

Corresponding author:
Ronald A. Beghetto, Department of Educational Psychology, University of Connecticut, Storrs, CT, USA.
Email: ronald.beghetto@uconn.edu

This is a wicked problem that we are talking about today, the idea of creativity, personalized learning, and large-scale assessments (LSAs) coexisting together. What I want to highlight are some paradoxes among these constructs and some possibilities. My background is in creativity studies. I am not a psychometrician, but I would like to take some time to explore the question of whether personalized learning and creativity can operate in the context of LSA formats.

I'll start by highlighting a few working definitions and my assumptions going into this talk. I will then highlight some of the paradoxes that result from considering these concepts together. Next, I will highlight one way to think about the logic behind LSAs and what it would need to look like to accommodate creativity and personalized learning. I will then provide a couple of examples from the Programme for International Assessment (PISA) 2012 Creative[1] Problem Solving Assessment and discuss whether it is appropriate to describe it as a measure of creative problem-solving. I will close by highlighting some possibilities and speculations about LSAs in relationship to creativity and personalized learning.

Definitions and assumptions

LSAs

Everyone attending this talk likely has a good understanding of what is meant by LSAs. In the context of this talk, I refer to LSAs in education as a standardized measure of some construct (typically academic achievement) in a large, representative sample of interest (e.g., Secondary students). LSAs are used to make generalized inferences about the broader target population or make comparisons across populations. PISA's 2012 Creative Problem Solving Assessment is an example of a LSA.

Creativity

So what do we mean by creativity? Creativity has been defined as having two key criteria: originality, newness, difference, or uniqueness *and* effectiveness, meaningfulness, or meeting task constraints (Plucker, Beghetto, & Dow, 2004; Runco & Jaeger, 2012). What I boil this down to for the purpose of this talk is any new, original, or different way of meeting some goal or set of criteria at the individual or social level. Creativity at the individual or intrapsychological level has been called *mini-c* creativity (Beghetto & Kaufman, 2007).

Mini-c creativity refers to the personal experience of creativity when learning something new and meaningful or having a creative experience. Creativity, of course, can also be recognized at the

social or interpsychological level by external audiences. When creativity is recognized by others, it is often referred to as larger-C creativity (Kaufman & Beghetto, 2009), which can range from everyday or *little-c* creative expression (e.g., a student comes up with a unique way of solving a problem in the classroom) to professional or *Pro-C* creativity (e.g., a scientist publishes a paper that contributes a new insight to the field) and historically recognized creative accomplishments or *Big-C* creativity (e.g., Madame Curie's work on radioactivity).

In educational contexts, we are typically interested whether mini-c creative ideas and insights can be shared, recognized, and potentially make a little-c creative contribution to others. So, for instance, can students identify their own problems to solve and their own unique way of solving them?

Personalized learning

Next, there is personalized learning. The way I define it in this talk is simply an idiosyncratic learning process. That is, personalized learning refers to a unique, personal, or individual process that results in new knowledge, understanding, skills, or behaviors. Personalized learning can also include what students decide to learn. In this way, personalized learning can be thought of as a form of creative learning because it is an idiosyncratic process that results in new and meaningful learning for the individual. More specifically, the path of learning is a creative path.

In fact, in the field of creativity studies, some of the earliest scholars recognized that we really cannot have an adequate theory of individual learning, unless we somehow account for the role creativity plays in the process (Guilford, 1950, 1967). In this way, learning is a creative act. This is a long-standing assertion. I think if we look at the definitions of creativity and personalized learning together, we can clearly see the connection. What I will be doing in this talk is assuming that personalized learning is thereby subsumed under creativity.

These are my operating definitions. Now, let's turn to paradoxes that emerge when we try to think about the combination of creativity, personalized learning, and LSAs.

Paradoxes

When we consider the features of LSAs, personalized learning, and creativity (as displayed in Table 1), the paradox of trying to combine them together becomes evident.

With respect to LSAs, the emphasis is on sameness. This emphasis on sameness, by the way, also tends to be found in K12 schools. We typically group students of the same age, doing the same thing, in the same way, at the same time, trying to obtain the same outcome (Beghetto, 2019). Sameness in LSAs is reflected in the fact that LSAs tend to be standardized measures. Test designers want to control for or remove any interfering factors that could result in inaccurate inferences about observed differences in scores between test takers.

Table 1. Features of large-scale assessments, personalized learning, and creativity.

Large-scale assessments	Personalized learning	Creativity
Sameness	Difference	Difference
Predetermined	To-be-determined	Retrospective and to-be-determined
Static	Dynamic	Dynamic
Generalizable	Idiosyncratic	Situated

Consequently, successful performance on LSAs pertains to completing the same (or equivalent) problems, under the same testing conditions, which have the same predetermined answers. How students arrive at those answers is typically not assessed or considered relevant. Now, what becomes paradoxical when we start looking at the features of these concepts together is: Personalized learning is about difference—sometimes profound individual differences—in students' own unique path to understanding something. This does not mean that criteria or standard approaches do not apply, but rather that students' idiosyncratic learning trajectories influence their current understanding and future learning trajectories (Schuh, 2017). Creativity, of course, is also about difference. It requires meeting goals or some set of criteria in different ways. It also involves identifying new problems to solve. So already we have this opening paradox between creativity and personalized learning in the context of an LSA.

Almost everything on the LSA is also predetermined. Almost everything is specified in advance. Conversely, personalized learning is *to-be-determined* . . . determined by the learner. Students' pathways are determined by their own idiosyncratic interests, needs, and aspirations. This is not to say that students don't need, wouldn't benefit from, or should not receive support from educators. Rather, the unique pathways of personalized learning often entail unexpected aspects for both the learner and the educator trying to support that learner. And of course, creativity is to-be-determined. Creativity is something that is emergent, and, in fact, I feel the best way that we can talk about creativity is as a retrospective distinction that we bestow on a variety of phenomena.

Creativity is determined after-the-fact (based on the criteria mentioned earlier), we say "that was a creative idea," "that was a creative behavior," or "this is a creative artifact." But in the process of being developed, we do not know with any level of certainty whether it will be judged to be creative. There are no guarantees. We cannot sit down and say, "I'm going to create a painting that everyone in this room is going to think is creative." That comes after the fact. It is a judgment made based on the criteria that we talked about: newness and effectiveness, or in this case, a different way of meeting outcomes. So again, we see the connection between personalized learning and creativity but it's at odds with LSAs and oftentimes also at odds with the ways we judge success in schools because those too are predetermined.

LSAs also tend to be static because of the outcomes they produce. What I mean by this is: LSAs tend to produce one indicator that is fixed and final—you get your score, your test score. It is by design summative. You may be able to retake an LSA and get a different score. But it is still designed to be summative and fixed, rather than open and inconclusive like creativity or personalized learning.

Personalized learning, in the way I conceptualize it, is much more dynamic. It's ongoing. Personalized learning is formative, it continues taking shape. Even if you attain your goals, they transform into new goals. So it doesn't make sense to say that personalized learning is over and the same can be said about creativity. Creativity is this dynamic and emergent phenomenon. It is a process and a product. So even when there is a creative product, the judgment about it being creative is still open to new interpretations (Beghetto & Corazza, 2019). Indeed, an idea once considered creative may later be considered banal and an idea once considered irrelevant may later be considered creative.

If you think about it, you can likely identify several examples where something was not considered creative and only after the fact you realize, "My goodness. This was a new and meaningful way of doing this . . . " and there are historical examples where people only later recognize that some contribution changed an entire field be it in the arts or sciences. Or it might be the case that we judge something in the moment to be creative and later recognize, "Having thought about it, I now realize that it isn't that creative." Judgments of creative outcomes are dynamic and thereby tend to be at odds with the summative outcomes of LSAs.

The other component here is LSAs are designed to be generalizable, so you can make inferences about broader populations. Whereas personalized learning, as I have mentioned, is idiosyncratic. And judgments about creativity are situated both temporally (in a particular time) and contextually (in a particular place). So what is determined as creative in a fourth-grade classroom may not be determined as creative in a different fourth-grade classroom. Or in an eighth-grade classroom or next year. So, creativity is very dynamic and situated.

Now, there are, of course, examples of creativity that stand the test of time. As mentioned, we call that Big-C creativity. But the creativity I am talking about in educational contexts tends to be rather ephemeral. A student's unique contribution to a class discussion may be forgotten, unless it really impacts someone's way of thinking or somebody documents it. Otherwise, it's gone. So these are some of the paradoxes.

Next, I would like to highlight these ideas in the form of some logic models, which may help clarify how the logic of LSAs tends to seal out opportunities for creative learning and what kinds of openings are necessary for creative expression.

Logic models of LSAs and creative expression

Let's consider the logic of LSAs, which also tends to apply to school-based learning experiences (see Panel A, Figure 1). As illustrated in Panel A of Figure 1, all the features have been
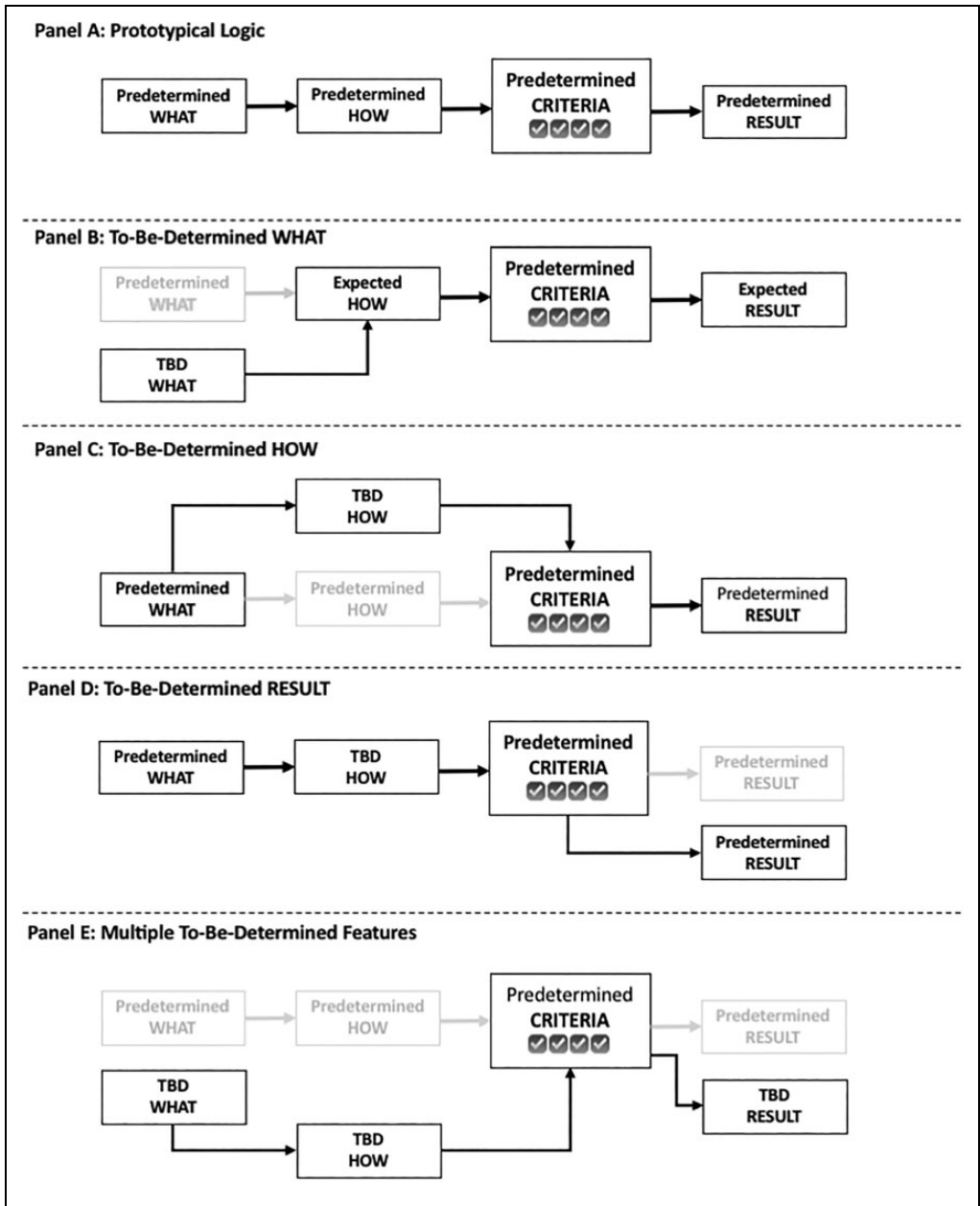
Figure 1. Protocol logic model and variations necessary for creative learning.

predetermined. LSAs, like many school-based learning experiences, predetermine *what* tasks students are expected to complete, *how* students are expected to complete those tasks, the *criteria* for successful performance, and the expected *result* (Beghetto, 2018).

All the problems have been specified in advance. There is a bit more flexibility in how students can solve a problem, but typically there is a rather clear expectation as to how students will work through a particular problem using a standard process or heuristic to meet the predetermined criteria and yield a predetermined result. Almost everything is known in advance.

As I mentioned, this logic model also serves as the basis for how we typically design lessons in schools. So there is this correspondence between how we design assessments and how we design the learning experiences in classrooms. This is fine if we want to assess whether students can accurately reproduce what has been taught. When it comes to creative learning, however, I would argue that the problem with this model is that it effectively engineers out meaningful uncertainty (Beghetto, 2018). If different pathways are not allowed, then there is little room for creative expression.

In school-based learning experiences, we often do not give students opportunities to come up with their own problems to solve and this is certainly the case on LSAs. We also tend not to acknowledge or reward different pathways for meeting the criteria. So students tend to get rewarded on LSAs (and also in classrooms) for meeting predetermined criteria by producing what is expected and how it is expected.

What if we started thinking about this logic model a bit differently and started opening things up? What if we allowed young people to specify a different problem (i.e., come up with their own WHAT) and maybe still use an expected HOW to meet criteria and yield an expected RESULT? Doing so is what I call lesson unplanning (Beghetto, 2018) or, in the case of the prototypical logic model, *component unplanning* (see Figure 1, Panel B).

If a different WHAT is allowed to be determined by students, then we introduce some structured uncertainty and thereby an opening for potential creative expression. Specifically, the uncertainty is still structured by the predetermined criteria to be met, the predetermined process, and the predetermined result. But, here students have an opportunity to do something new and different that maybe educators or test designers didn't even realize. That is one permutation.

Another variation would be allowing for a different HOW (see Figure 1, Panel C). This provides students with different ways of solving their problems, and in truth, many assessments already do allow test takers to solve problems in whatever way they want. In most cases, however, students are not scored on how they solved the problem. So students do not receive feedback or acknowledgement for coming up with a different, yet accurate (i.e., creative) way of solving problems. As we will see when I briefly discuss the PISA 2012 Creative Problem Solving Assessment that the exam does open up the "how" and even encourage exploration of the how, but students are not rewarded for or given extra points for original or unexpected pathways. In order for assessments to provide feedback on and to assess creative expression, they would have to assess the originality

and meaningfulness in how students solve problems (not simply whether students use a predetermined process to arrive at the correct answer).

Looking at Panel D of Figure 1, the RESULTS can also be opened up, which would require specifying the kinds of problems that can lead to multiple outcomes and different solutions. And finally, we could really open it up and have students specify their own problems, their own ways of solving those problems, which would result in different outcomes, but still meet the predetermined CRITERIA (see Figure 1, Panel E). If you really want to get radical, maybe even allow kids to define their own criteria for success, which would be the far end of personalized learning and creative expression continuum.

Examples from PISA 2012 "Creative" Problem Solving Assessment

We have been getting a bit theoretical and conceptual here. So, let's now situate our discussion in some concrete examples from PISA 2012 "Creative" Problem Solving Assessment. I put "creative" in quotes because I want to interrogate whether these might be considered *creative* problem-solving items or are they just simply interactive problems that have some ill-defined components, but not really used in a way to assess creativity? We will find out. You can judge for yourself. I want to show you a couple of PISA problems, which have been released online (see Figure 2).

So if you go online[2], you can work these problems yourselves. These are online delivered problems. They are dynamic. You get to click on things. You get manipulate different things and then you can submit your choices and see what you get. So you can work through the process and follow along as I describe them. In the interest of time, I'm going to just look at two examples (see Organisation for Economic Co-operation and Development [OECD], 2014, for a more thorough overview of the different types of problems used on the assessment).

Let's look at this first problem (see Figure 2, Panel A). This problem involves having young people solve a traffic problem. It's a predetermined problem. The problem is about two people who live in two different towns and they want to meet in a suburb on the map. No one wants to travel more than 15 min. Where could they meet? The process for solving it is somewhat open-ended.

The criteria are clearly specified. The problem is clearly specified and the solutions are predetermined. All the necessary information is available and thereby represents the lowest level of problem-solving proficiency (see OECD, 2014). Even in this low level problem, the *how* is still somewhat open (see Figure 1, Panel C). There are two predetermined, but different solutions that PISA will accept on this problem. Students only need one to correctly solve the problem. How the kids do this is up to them, which I would argue is typical in most assessments. So is this creative problem solving? In one, somewhat narrow sense, we could say *yes*. It does afford a personal minic experience of this problem.
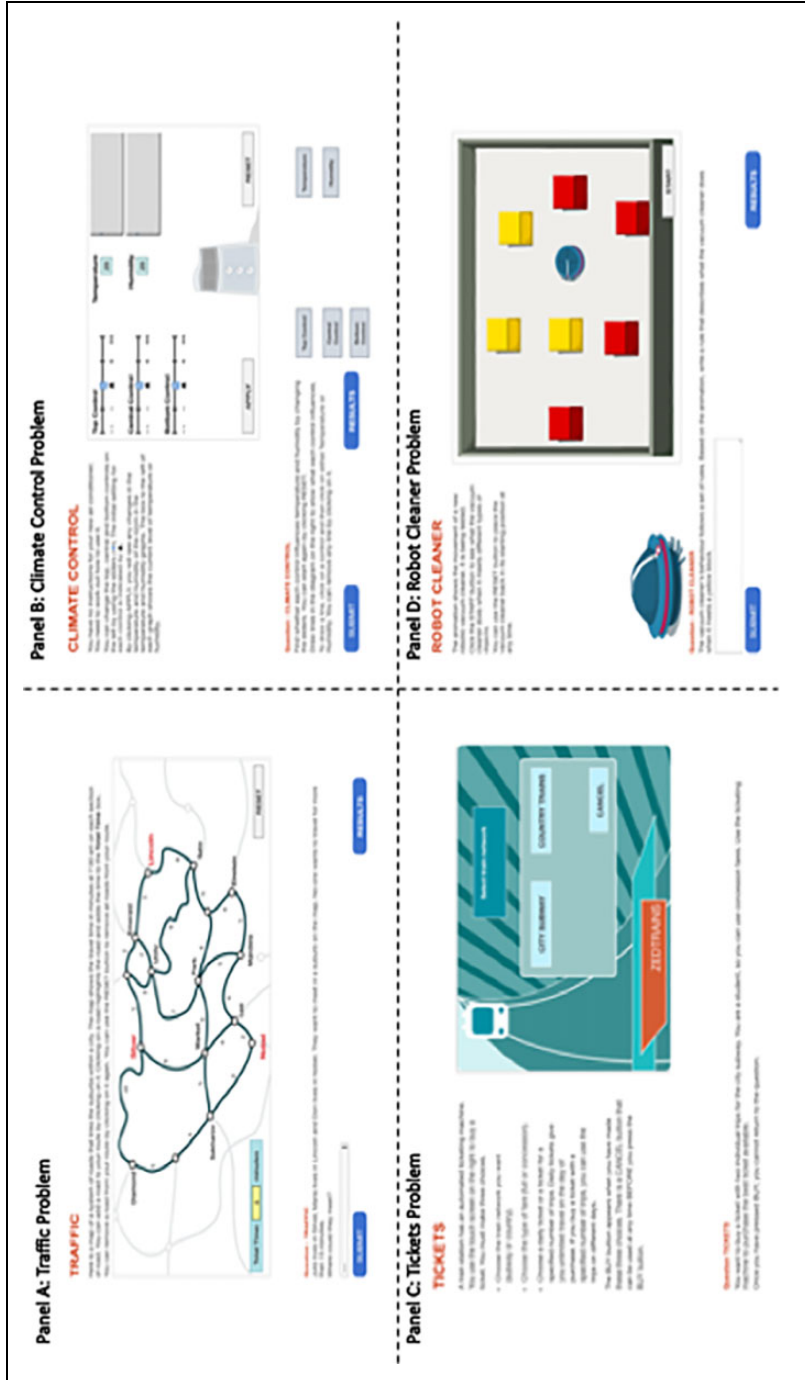
Figure 2. Example problems from PISA 2012. *Note:* Screenshots from this figure were retrieved from CITE. PISA = Programme for International Student Assessment.

319

A student tackling this problem, could say, "Wow. What I just did is a new and meaningful way of solving this problem. It's new and different to me. It's creative to me." But the test designer or teacher might say, "Yeah. That's a pretty common way of solving that problem." Being new to the student but nobody else is mini-c creativity. So the student theoretically could have a mini-c creative experience while working on this problem. But, the problem itself is not assessing creativity in problem-solving, because the originality of the process is not assessed and no points are awarded for more or less original ways of solving the problem. Students are only assessed based on whether they arrive at one of the two predetermined correct answers.

Generally speaking, when students work through problems on any assessment, it may be possible that they generate a solution that is quite unique. If you were able to look how each test taker responded, then it is possible that some students solved problems in original ways—ways that others would recognize as creative (i.e., little-c creativity). That would be interesting. That's what we want to hope for in an assessment of creative problem-solving. We don't want to restrict assessment experiences to mini-c (i.e., only being new to the student). We want real uncertainty in the problem so students can generate creative responses recognized by others. But, again, the process would need to be assessed or judged for creativity.

Let's look at quick example of another problem (see Figure 2, Panel B). This one has a little bit more going on with the problem. Again, the *what* is specified, the *how* is up to the student, the criteria are specified, and the outcomes are, of course, predetermined. What the test taker has been asked to do here is to figure out the controls for this air-conditioning system. There are three sliding controls on it that they can manipulate (labeled: *top control, central control, and bottom control*). Changes in the sliders may change the temperature and humidity. You will see a little line chart and numerical indicator for temperature and humidity change if you move a slider that influences temperature, humidity, or both. You can reset it, keep trying different things. Ultimately, what the test takers are trying to do is to figure out which of the three sliders control temperature and humidity. Does the top control influence temperature and humidity? Does the center control influence humidity or temperature? and so on.

Once test takers figure it out, they have to draw an arrow from each labeled control to temperature and/or humidity to record their answer. So this problem has a to-be-determined and somewhat ill-defined process for solving the problem. It is a problem that students have to play around with to solve. This problem affords a different HOW for students (see Figure 1, Panel C), but there is only one predetermined solution. So, same sort of thing as the previous example. A student may experience mini-c creativity or even come up with a little-c process for solving it, but we would never know on this test because the originality of process is not assessed or scored. Although, according to OECD (2014, p. 50, emphasis added),

partial credit on this particular version of the problem is given if the student explores the relationships between variables in an *efficient* way, by varying only one input at a time, but fails to correctly represent them in a diagram.

So, do these examples assess or provide space for creativity and personalized learning? The most generous response would be: *not really*. Not much more so than most other LSAs. PISA 2012 Creative Problem Solving Assessment does offer interactive experiences and even somewhat ill-defined problems, which test takers may feel is different from what they are used to solving on academically oriented LSAs (see OCED, 2014 for details). Consequently, it is possible for test takers to have a mini-c creative experience when using their own personal way of solving the problems (i.e., have a new and personally meaningful insight during the problem solving process), but larger-C creative and idiosyncratic approaches are not scored by the assessment.

Students get scored on whether they solve the predetermined problems by providing the pre-determined solution. Some of the items are scored for how students solve them. But, again, the assessment does not assess originality or creativity of responses, just whether students do what is expected and, in cases where credit is given for process, how it is expected.

Possibilities

Let's now talk about some possibilities. We will return to the PISA 2012 examples, but we'll also push beyond them. So what we can do to open things up a bit to assess creative expression? I will offer three possibilities and speculations about what we can do to think about it. One is collecting and using process data from LSAs to explore how students are working on a problem. If you collect that type of data, you can analyze it at the individual student level. The good news is PISA actually captured some log-file process data and has released it for analysis.[3] The log-file data capture clicks of individual students who take the assessment. The data is available as raw data for analysis.

The other thing you could do is take items that are released and use them as stimuli to study how students work through such problems. So you can do different things with LSA items and data to potentially examine creativity, even though the assessments themselves do not assess creative learning. You could assess released data sets by designing small-scale studies of LSA items or you can design smarter LSAs. Let's briefly take a closer look at each of these options.

Analysis of released data sets

First, let's consider what might be possible using existing student-level log data from PISA 2012 Problem Solving. PISA collected student-level process data on what test takers clicked on, duration of solving the problem, different events, how students manipulated the controls in Climate Control problem (see Figure 2, Panel 2), start and finish times, whether they arrived at correct responses,

and so on. You also have their overall problem-solving score as well as any other data collected from PISA surveys.

But when analyzing extant data, you need theory. You need to situate these kinds of analyses in a relevant theory. You need an understanding of creative problem-solving to analyze these data to make any sense of the question of whether the process data can be considered creative. So analyzing the *how* (i.e., the process) in relation to the solutions obtained, one possibility would be to score the approach for originality (i.e., the statistical improbability or evaluated uniqueness of an approach compared to other approaches in the sample, see Carruthers & MacLean, 2019; Kaufman, Plucker, & Baer, 2008). You can look at how different a student's pathway is compared to all the other pathways in the data set. And if the response is both original *and* accurate, it would thereby be considered creative. So you could actually examine variations in creativity scores based on levels of originality and accuracy.

Here's an example of a study that used log data from the PISA 2012 Climate Control Problem (Greiff, Wüstenberg, & Avvisati, 2015). This team of researchers looked at the log-file data. The nice thing about this and the reason why I think it is a good example is: The researchers started with a middle range theory. They take this interesting data set and use a theory of problem-solving, not necessarily creative problem-solving, but problem-solving called the "vary-one-thing-at-a-time" (VOTAT) strategy. In this approach, you hold everything else constant as you change one thing at a time. The researchers had a prediction based on this problem-solving theory. And they found by looking at the PISA process data from this problem that kids who employed the VOTAT strategy had, on average, higher performance on this problem and on their overall problem-solving score.

This study represents an interesting possibility. It highlights how it can be possible to design a new study based on the analysis of raw data collected by PISA, which was not initially used to assess or score *creative* problem-solving. On the PISA exam, the focus is really on whether test takers solve the problem accurately and, on some problems, whether they use predetermined approaches to solve those problems. So, for the purpose of assessing creativity, it is theoretically possible to use log data to assess originality of responses, but it would require secondary analysis of the released data guided by a relevant theoretical framework of creative expression.

Smaller scale studies of LSA items

Another way you could potentially explore creative and personalized learning is to use released items from LSAs and design more detailed process and outcome studies. But again you need theory to do so. If you just analyze data without a theory and try to identify and describe meaningful patterns, you're just engaging in at best rigorous journalism and, at worst, blind empiricism. Theory allows you to consider and test possibilities, which may or may not be represented in the data set, and there are plenty of robust theories from the field of creativity studies that you can draw

on. Researchers can also draw on theories to guide studies designed to explore individual differences and personalized learning in relation to released LSA items.

Here's an example of how researchers (Selleri & Carugati, 2018) used theory and a released LSA problem from PISA 2012 assessment of mathematical literacy to design a study aimed at looking at different ways students approach and experience problems. They used the Mount Fuji problem and designed a qualitative study, which involved having teams of students working on this problem. They were guided by a theory focused on sociopsychological didactics of mathematics. So they were using a standardized test item as means for a microanalysis of how different students approach the same problem. They demonstrate how this kind of microanalysis can yield insights beyond what the test scores provide. Specifically, their work indicates how "final solutions could hide different difficulties of understanding the problems, devising a plan, carrying it out, managing mathematical procedures, and referring back to them" (p. 502).

This example highlights how problems used on standardized exams can be used as a means for designing studies that explore sociopsychological and more personalized approaches that students employ when engaging with standardized LSA items. Going even further, you can design studies that open up the design of existing problems (see Figure 1), so students can come up with their own questions and problems and their own ways to solving them. And even their own criteria for what would be a successful outcome and whether they can generate multiple outcomes that meet the criteria.

So another way of saying this is you could take the problem we saw earlier, the traffic problem, and you could say, "Alright, some of the constraints are predetermined, but you can open up the rest to come up with your own problem? How many different ways can it be solved? How many different solutions are possible?" So more fixed LSA problems can be transformed from predetermined to to-be-determined. Doing so introduces the uncertainty necessary for creative expression, not only for the students but also for educators and researchers because we are not certain what students will come up with or do. That would be a genuine ill-defined, to-be-determined problem. This, in turn, increases the chances of seeing creative expression.

You can then go further and overlay other micro-longitudinal features, which is what some of my colleagues and I have been working on, to explore not only creative expression but other related and dynamic constructs such as creative confidence and emotional states during problem-solving (e.g., Karwowski, Han, & Beghetto, 2019). Using multiple, brief intervals of measurement, researchers can more dynamically measure features and related factors of creative performance on tasks (e.g., creative confidence, confidence in completing the task, emotional states, and so on).

A kid who is very confident about creatively solving problems in general might feel hardly confident at all when presented with a particular problem in a specific context. Another student who starts out with low confidence can have that confidence grow over the course of successfully

working through a task. The point is, we likely will see variations in creative confidence between and within individuals. Exploring these different profiles of creative confidence and other relevant variables can yield important insights into the nature of creative problem-solving.

You can also have a panel of judges assess the creativity of the process or completed task. This kind of thing goes well beyond the typical black box of a correct or incorrect score on LSAs. This is an example of where we would need to be moving toward to actually get into assessing and understanding individualized learning trajectories and creative problem-solving. In the field of creativity studies, researchers are starting to do this kind of work which allows us to consider patterns of within and between variance in how students experience and perform on creative tasks.

I want to briefly return to what I mentioned about using judges to assess the creativity of a product or process. Teresa Amabile, a social psychologist from Harvard Business School, developed an approach (Amabile, 1996), which she calls the consensual assessment technique (CAT). It is a very popular approach in the field of creativity studies because it can be used to assess the creativity of a solution to a problem and it can also be applied to assessing the creativity of a problem-solving process. You get experts in the domain to independently judge creativity. If it is poetry of students you want to judge, you will get published poets. If it is performance on a math problem, you get mathematicians. The validity lives in the expertise of the judges. You have them individually assess what students have produced and calculate the inter-rater reliability among the experts.

If you have obtained five or six experts, you look at the reliability. If you have acceptable reliability and you have the validity of the expertise, then you have an assessment of creative expression. The judges are comparing creative performance on a task in relation to how others performed on that task (see Baer & McKool, 2009 for an overview). The comparison is made within a particular performance setting, a test taker's creativity compared to other kids' performance.

So, there are ways of assessing creativity and personalized approaches to solving test items using these preexisting items, but this is not what LSAs are typically designed to do. PISA 2012 Creative Problem Solving does have an item that requires expert scoring of student responses (see OECD, 2014), but again it is not scored for creativity. Expert scoring of creativity requires an entirely different approach and it can be very time- and labor-intensive, which leads us to the final possibility or speculation I want to highlight.

Designing smarter LSA's

I am going to close with the possibility of designing smarter LSAs and highlight three points. First, if we really want to assess creative learning at the individual level, we have to present students with a range of complex tasks. What I mean by complex task is increasing uncertainty (not just

difficulty). I'm referring to real uncertainty—where it's not just ill-defined for the kid, there is uncertainty for everyone in how students are actually going to meet the criteria. So if there is low uncertainty, that represents a simple task. It does not mean it is easy. It just means it has low uncertainty. Conversely, if we have a task with multiple features that are to-be-determined, then we would have a complex task.

Second, we need to capture additional process and experiential data while students are working on problems. And finally, it is worth considering and speculating about possibilities for human-artificial intelligence (AI) augmented scoring. We could, perhaps, use machine learning to train on a data set of expert-scored creativity (e.g., based on Amabile's CAT approach described above). We are in a time now where machine learning algorithms could likely be developed that could learn from and eventually match the expert ratings of creativity from human judges. This would address the somewhat prohibitive time and effort needed to impanel a set of human experts to judge each student's response to items. Machine learning and AI may be able to address this challenge and perhaps even result in LSAs that assess creativity using approaches currently used by creativity researchers and perhaps lead to new ways of assessing creativity.

So just to give you another way of looking at these three points. A continuum of complex tasks can look something like this. On the current PISA-designed creative problem-solving tasks, the problems provide extremely limited opportunities for creative expression because the only thing that is to-be-determined is the HOW, which is not even scored for creativity on the assessment. So, at the very least, that would need to be scored for creativity to be considered an assessment of creative problem-solving.

Even if that happened, we would still be limited to moderately complex tasks (with respect to the uncertainty involved) because you really are only allowing for different ways of solving the problems. We would need to include more complex tasks that allow students to also develop their own problems, their own ways of solving those problems, multiple outcomes for problems, and potentially even their own criteria for success. That would provide a more robust continuum of complexity necessary for assessing creative learning and problem-solving.

As far as capturing different forms of experiential data, there are a variety of different ways that you can assess experiential data. Researchers have an increasing variety of ways to measure in-the-moment behaviors, emotional responses, physiological states, confidence beliefs, and so on. We need multiple types of measures and more robust data if we are interested in looking at how students actually approach and experience problems in their own unique way.

Finally, machine-learning approaches trained on how experts have judged the creative process and products of particular problems seems like a more and more likely possibility. But it would require more than just assessing originality. Creative responses involve assessing unique and

different ways of meeting task constraints. Recall too that judgments of creativity are situated so they may not generalize in ways that results of LSAs are typically used.

In closing, I would say that perhaps it is possible to design more sensitive and smarter LSAs, which could assess creative and personalized learning. But as far as I know, we are not there yet. And even if we can get there in the near future, it is critically important for us to pause and seriously consider whether we should. What are the potential costs and benefits to young people, educators, schools, and societies?

Author's note

This article is based on an edited transcript from an invited talk given on October 28th, 2018, at East China Normal University's Forum on Educational Research.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. PISA sometimes refers to this assessment as "Creative Problem Solving" (https://www.oecd-ilibrary.org/education/pisa-2012-results-skills-for-life-volume-v_9789264208070-en) and other times just "Problem-Solving" (http://www.oecd.org/pisa/pisaproducts/pisa2012problemsolvingquestions.htm).
2. The two example problems discussed in this talk: Traffic, Figure 2, Panel A, retrieved from http://www.oecd.org/pisa/test-2012/testquestions/question1/ and Climate Control, Figure 2, Panel B, retrieved from http://www.oecd.org/pisa/test-2012/testquestions/question3/. For a more detailed discussion of the different kinds of problems and what aspects of problem solving and level of proficiency they measure refer to https://read.oecd-ilibrary.org/education/pisa-2012-results-skills-for-life-volume-v_9789264208070-en#page1
3. Available here http://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm under "Log-file databases for released PISA 2012 computer-based items."

References

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.

Baer, J., & McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In C. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65–77). Hershey, PA: IGI Global.

Beghetto, R. A. (2018). *What if? Building students' problem-solving skills through complex challenges*. Alexandria, VA: ASCD.

Beghetto, R. A. (2019). Creativity in classrooms. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 587–606). New York, NY: Cambridge University Press.

Beghetto, R. A., & Corazza, G. E. (2019). *Dynamic perspectives on creativity: New directions for theory, research, and practice*. Cham, Switzerland: Springer Nature.

Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *1*, 73–79.

Carruthers, L., & MacLean, R. (2019). The dynamic definition of creativity: Implications for creativity assessment. In R. A. Beghetto & G. E. Corazza (Eds.), *Dynamic perspectives on creativity* (pp. 207–223). Cham, Switzerland: Springer Nature. Retrieved from http://link.springer.com/10.1007/978-3-319-99163-4_12

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, *91*, 92–105.

Guilford, J. P. (1950). Creativity. *American Psychologist*, *14*, 469–479.

Guilford, J. P. (1967). *Creativity and learning*. Oakland: University of California Press.

Karwowski, M., Han, M., & Beghetto, R. A. (2019). Toward dynamizing the measurement of creative confidence beliefs. *Psychology of Aesthetics, Creativity, and the Arts*, *13*, 193–202.

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

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York, NY: Wiley.

Organisation for Economic Co-operation and Development. (2014). *PISA 2012 results: Creative problem solving. Vol. V: Students' skills in tackling real-life problems*. Paris, France: OECD.

Plucker, J., Beghetto, R. A., & Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*, 83–96.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*, 92–96.

Schuh, K. L. (2017). *Making meaning by making connections*. Dordrecht, The Netherlands: Springer Nature.

Selleri, P., & Carugati, F. (2018). Errare humanum est! A socio-psychological approach to a "Climbing Mount Fuji" PISA question. *European Journal of Psychology of Education*, *33*, 489–504.