



Ryan Cody Hatcher

University of Connecticut

Introduction

Recently creativity has become a focus of study in the educational community because of its potential influence on motivation, self-efficacy, innovation and general well-being. Moreover, in the study of creativity the ability to reliably assess creativity has become an important research topic. Many believe that if we could assess creativity in an effective way we could identify an individual's creative ability, an individual's creative potential in a specific or general domain, and influence how we facilitate development in those domains. The problem, however, is there is no reliable method for assessing creativity. Right now the favored method is the consensual assessment technique (CAT), but it is difficult and time consuming to gather experts to do ratings. The CAT is only one method of assessing creativity but all the other favored methods today suffer from similar or worse problems. It is a very serious concern and an area of great opportunity for development. Everyone can agree that reliable creative assessment would be a noble goal but those that have the most to gain include teachers for students, researchers for subjects, curriculum planners for school curricula, colleges for admissions, business managers for employees, and individuals for themselves. To address this challenge and opportunity I propose developing a web-based creativity rating platform situated, at first, in a specific community of practice that would allow creativity assessment and would be assessable by all users in that community.

WebCAT

The webCAT stands for web-based consensual assessment technique. It is a web, cloud, or cell based platform from which users act as the raters to assess the creativity of a specific product (or person). As a web-based tool I envision it being accessible from a website that serves a specific community of practice, at first. Eventually I want it to be used by everyone with access to the internet either in real-time or post-activity. In developing this tool I hope to address some of the concerns listed earlier, specifically the problem of developing a reliable creativity assessing tool. By situating it first within a community of practice, the practitioners would act as the experts to assess creativity using the CAT. It will also allow non-practitioners access to a large pool of experts within a specific domain. In addition, as an added benefit, I believe it will also serve as an effective method of developing rater creative meta-cognition, as well as, helping a community converge on a general but dynamic understanding of the definition of creativity. My reasoning and justification for specific design decisions are based upon work already done in the field of creative assessment and web-based assessment tools. There has not been much in this area but I have made an attempt to analyze a portion of it.

Literature Review

The use of assessment in creativity has been a short but frenzied endeavor, which has seen the creation of a large number of instruments to measure a wide range of factors of and variables that relate to creativity. In the project I am proposing I have made use of the consensual assessment technique (CAT). However, it is by no means the only instrument I have looked at and indeed it may not even be the most suitable for creativity assessment in a digital / virtual space. In this review I outline some of the design decisions I have made as well as what advances into the digital / virtual space have already taken place within the field of creativity assessment.

When designing a program or assessment instrument for any product bias is an important consideration. It may be nearly impossible to escape completely but with due consideration and thoughtful design work bias in assessment can be mitigated. In thinking about bias there have been many studies that have showed how it can influence assessment; one such example is from Lebuda and Karwowski (2013) who conducted a study looking at name and gender bias in creativity assessment using a modified version of the CAT. The researchers found that there was a statistically significant link between name, gender, and assessment. They found that a unique name created a more positive perception in the domains of music and poetry. A male name generally scored higher overall. Female names received especially low ratings in the domain of science. Finally, they found that highest creativity score was given to females with unique names and the lowest scores were given to men with unique names. Given the limitations in the study related to the number of expert raters and statistical significance found during comparisons it is difficult to arrive at a conclusive cause and effect relationship; however, the implication implicit in the results present an extremely interesting biased interaction and one which should be given considerable attention in creativity assessment in the future.

Along with names, gender and racial bias can have an important influence on creativity assessment. A study by Agars, Baer, Kaufman, and Loomis (2010), examined whether or not providing expert raters the racial and gender information of an individual had a significant effect on assessment ratings using the CAT. From their study they found that their raters showed little to no bias. This confirms that the CAT is generally good at suppressing bias in ratings; however, it demonstrates that these biases can create some noticeable difference in creativity ratings and as such should be considered in any development of an assessment tool. To address the concerns raised in these two studies it is important then to build a safeguard into the design of any program seeking to assess creativity that all submissions should remain anonymous for the duration of a rating period.

These two articles taken together show the importance of both biases that can be influence creativity assessment –even just a little—and the importance of differentiating between expert raters for use in the CAT. Baer, Cole, Kaufman, and Sexton (2008) conducted just such a study to determine the importance of and differentiate between the different levels of experts used in the CAT. They found that expert raters created the most consistent ratings. They also found that non-experts (novices and intermediates) agreed less with the expert raters and less consistently with each other. In translating these results to the development of a creativity assessment tool the determination of what is an expert and what is not an expert is paramount in insuring high validity and reliability. Thus in designing a digital / virtual CAT tool, where there may be hundreds or thousands of potential raters, it becomes imperative that there is a system by which experts and non-experts can be differentiated.

Changing directions briefly I would like to look at instances in which creativity and assessment have been implemented in a digital / virtual environment and the viability of doing more robust research within these environments. A study conducted by Hass (2015) looked at the feasibility of using the divergent thinking (DT) test with web-based methods. Although, there is much debate about whether the DT test is actually a measure of creativity, the study presents an important step forward in determining the usefulness of adapting standard measures of creativity and intelligence into a digital/virtual space. Hass's study explored two questions, (1) if it is feasible to assess creativity—with a DT test—using web-based methods and (2) how the subjects would react to testing constraints unique to an online setting, such as time. In relation to the first question Hass found that testing of creativity using the DT and verbal fluency was indeed possible and reliable using web-based methods. In relation to the second question Hass found that time constraints did have an influence on participant responses, however, the most important factor in determining responses was task type not time limitations. This is an important first step in establishing the viability, usefulness, and adaptability of pre-existing creativity assessment

instruments and serves as an important resource for planning and designing more web-based tools in the future.

The next article is not a study but is instead a source of inspiration and ideas. In their article, Pretz and Link (2008) present a digital / virtual creativity tool, the creative task generator. This is not an assessment tool by which raters rate a creative product, but more a means by which to generate authentic creative products to be assessed. Their web-based tool is a program for the creation of divergent-thinking and open-ended creativity tasks and developed using several programming languages which can serve as a foundation for more digital / virtual instruments in the future. Also, in their tool they have built in a way for the creative products to be assessed using the CAT. Pretz and Links article and product present a significant step forward into online space and have helped to generate many ideas for potential integration of creative. The findings from Hass and the ideas from Pretz and Link together are very important in helping researchers think about more and better ways creativity assessment can be implemented.

All of these articles provide a great basis by which to begin thinking of how to design and implement a web-based creativity assessment tool. Specifically these articles make clear the challenges in establishing universal anonymity, the importance differentiating large and diverse populations into expert and non-expert raters with the CAT, and ways to successfully use existing creativity tools in a new environment as wells as potential new ways to address creativity assessment in a digital/virtual space.

WebCAT Model

As stated the WebCAT will be an online tool to be integrated onto a website that serves a specific community of practice, similar in design and function to metacritic.com or imdb.com, and

programmed in PHP, MySQL, and JQuery / Javascript. In this introductory scenario the website in questions will be Englipedia.net, a web-resource for ALT / ESL instructors. Englipedia is designed for users to share lesson plans which then become available –for free— for other users to view and rate, if so desired. The lessons are categorized into elementary school (ES), Junior high school –middle school- (JHS), and high school (HS). There are further categories on the website which include articles, blogs, forums, submissions, links, and warmup activities (not full lesson plans, but applicable to all levels. The program breakdown will be presented in this context exclusively—at this point—but some parts of the program may be generalizable.

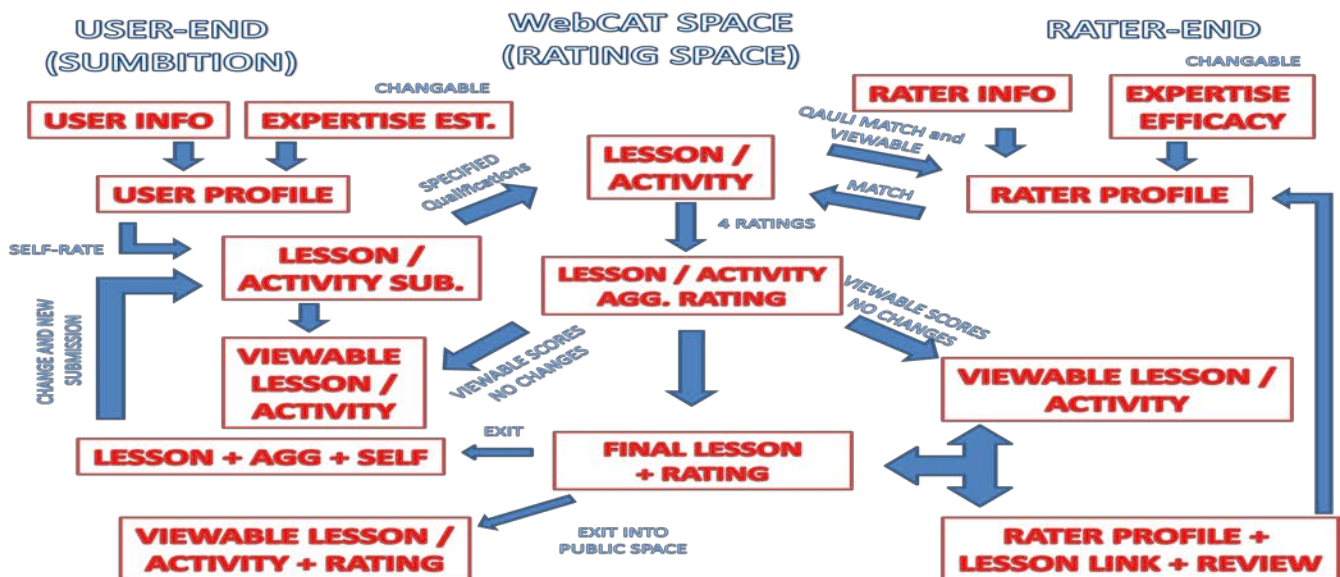


Figure 1. WebCAT flow diagram.

I will now seek to explain Figure 1, the parts and flow of the program. It is important to recognize that there are four distinct sections in this diagram: (1) the user space, (2) the rater space, (3) the webCAT or rating space, and (4) the public space. The program will directly influence and be influenced by the first three spaces. All outputs from the program will enter into public space but there is no input from public space into the flow of the program.

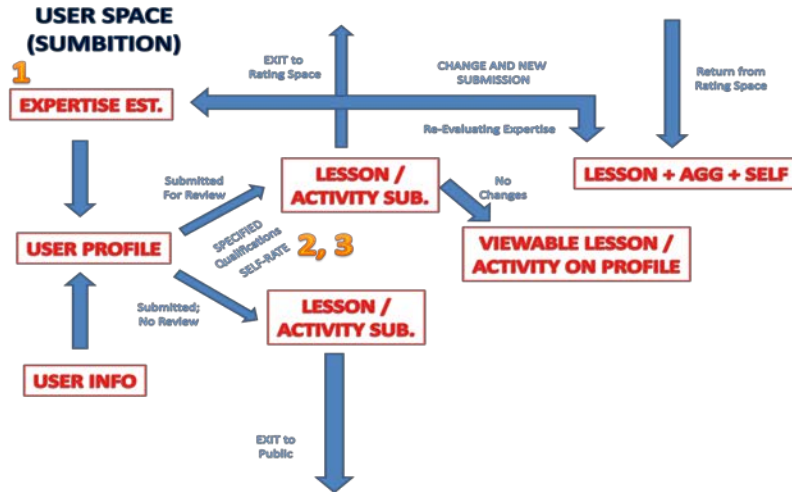


Figure 2. User space flow diagram.

In the user space the user will create a profile that is integrated into the profile creation system of the target site. In addition to the normal user creation steps the program will add an overlay to allow the user to estimate their expertise in the specific domain the user will be working in. The expertise rating will have four increments beginner, intermediate, advanced, and expert.

After creating a lesson, users can submit the lesson to be viewed in the public space. It is not necessary for users to seek a rating for their work if they do not want it rated. They may, however, select an option to allow the lesson to be rated—provided through the program—which will then move the lesson into the rating space. If a user chooses to have the lesson rated qualifications will be created for the lesson based on grade level, type of lesson, length of lesson, and level of expertise of raters requested. There is a question of whether to allow users to specify what level of expert raters they want; however, I believe that some raters may benefit more from having raters at a similar level or one step higher as opposed to only expert, expert raters. This coheres with some aspects of the sociocultural theory of learning. When providing scaffolding and making use of an individual's zone of proximal development (ZPD) it is better to use support that is only slightly more advanced than the user. After

submitting a lesson for rating the user has another option to self-rate the lesson using the same CAT criterion the raters will use. These three additions to the user space, the estimated expertise, the self-rating, and the specific expertise matching are being used to introduce an element of meta-cognitive development into the program. I hope to allow users to better self-regulate their own creativity with respect to the specific domain being employed.

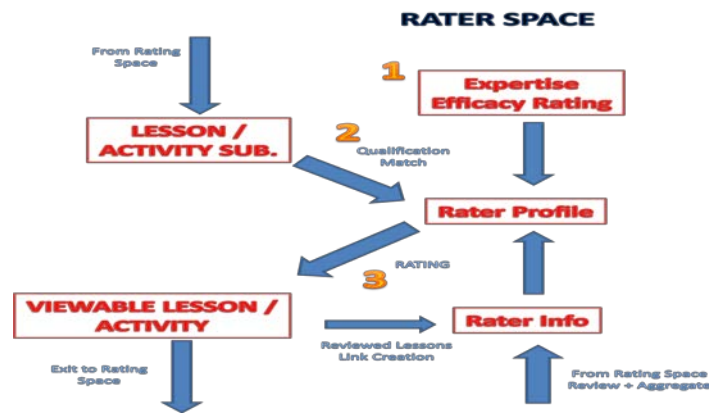


Figure 3. Rater space flow diagram.

The next space—the rater space—is similar in design to the user space. I had considered combining the user space and the rater space but I wanted to allow for a separate space for users to develop the aforementioned creative meta-cognition. A rater space would also allow for more flexibility in handling ratings and easier flow within the program. The rater will create a profile that is separate from the user profile. The rater will create an expertise rating that is based upon actual expertise in the specific lesson to be rated. To do this I have considered self-efficacy scales, surveys, or questionnaires. I have neither decided on a specific one nor developed questions to introduce to determine the rating; however, I recognize that whichever I use must be tested for validity and reliability. Additionally, creating an expertise rating must be simple enough that anyone can use it and short enough that raters do not find the process cumbersome. Along with an expertise rating raters must also create a list of

qualifications to allow matching from user submissions. Their qualifications are the same as the user designated ones, grade level, type of lesson, and length of lesson. After being matched with a lesson raters will rate the lesson using the same four factors that users will use in creating their self-rating—originality, usefulness, adaptability, and motivating. Additionally, they will then rank the four factors in order of importance in this specific domain. Once ratings have been completed the rating will be added to the aggregate calculations in the rating space.



Figure 4. Rating space flow diagram.

The final space this program will influence is the rating space. This space will be created and stored specifically on the program server. It will not be accessible to the public, user, or rater; it is a completely virtual space. It is responsible for matching the submitted lesson with individual raters via qualification matching. Once a rater has created a rating for a lesson the rating—not the lesson—is moved back into the rating space. After a certain threshold in number of raters is reached, the program will aggregate creativity scores by aggregating the scores for each factor, the ranking of the factors, and the total aggregate creativity score. At which point it will add the aggregate score to the user’s lesson in their profile, the rater’s profile, and to the lesson in the public space. The rating for an individual submission will be held in the rating space indefinitely. If a user changes a submission the program will

specify versions of the submission and create a new rating for the profile. In contrast a user can seek a rating from a different expertise level for the same submission, however, the rating displayed in the public space will be of the highest expertise level and raters at that level will not be able to see ratings done by lesser expertise levels.

Future Steps, Uses, and Directions

The future steps in developing this program are fairly straightforward. I will learn how to program in the previously listed programming languages. Next, I will attempt to integrate the program into a website that I have built for testing purposes. I want to stress that I will need to learn how to integrate the program into a website or program the tool as a layover for existing website coding so that the creativity assessment core program will be adaptable to specific features on a single website and changeable to assimilate to many websites. I very much fear that this will increase the time it takes to learn but it is a necessity to create something more accessible and easy to use with pre-existing assessment methods. After completing the programming I need to determine reliability and validity using many expertise levels as outlined in prior research dealing with the CAT. Along these lines I will also need to determine specifically how to accurately divide users into separate expertise levels. Additionally, I would like to do perception testing to determine if ordinary web users find the experience easy and intuitive.

After completion of the program there are many uses I can imagine for a digital / virtual creativity assessment tool beyond what I have already described. First, the program is designed to facilitate creativity in specific communities of practice and specific communities of learners, within education if you look at individual classes, schools, districts and even groupings of students in clubs, teams, etc. as a community of practice it is easy to begin to think of the myriad ways the webCAT can be used. For example imagine a creative writing class. The webCAT can be specially integrated into a

website that serves a community of practice composed of students from the class. Each student, teacher, and assistant is a potential user and rater. If you begin to bring in more classes the number of raters and the diversity in submissions will grow exponentially. If you take that same webCAT implementation and use it as a portfolio for the entire class you can begin to chart creative progress in individual students or the class as a whole over a long period of time. You can also create artifacts for parents and administration to review.

Another potential use for the webCAT is in college admissions. Bias is a very important topic in college admissions. As I outlined in the literature review the CAT and specifically my design and implementation of the CAT should minimize the role of bias in assessing students. Plus, it should not be difficult to integrate the webCAT into already existing online applications for colleges. Deans, professors, graduate students, and normal students can serve as expert raters for the likely huge diversity in creative products admissions are likely to receive. By open sourcing the raters it is likely that you will easily find raters who can fill the required expertise level.

I can also easily see the webCAT being a tool in business. In fact, I believe something similar to the webCAT already exists in specific business models that allow for open environments and lateral thinking in employees. Imagine if employees had more than one expert skill they can use to rate creative ideas (not an extreme stretch of the imagination I know); however, many businesses seem to want to limit creative development to the domain in which an employee was hired for. If business leaders could identify and leverage other creative domains in their employees it would be sure to be a boon for innovation and the bottom line.

As far as future directions for this project, I believe that conceptualizing the webCAT as a web-based iteration of the CAT is rather limiting. I believe that if the webCAT can be successfully integrated

into digital and virtual environments on a web-based platform it should also be possible and greatly desirable to consider other digital platforms. Along this line of thought I have considered developing the webCAT as a stand-alone program for PCs and Macs to assist users in developing community based websites and creativity databases and helping people to connect with like-minded and creative individuals. And, as I truly want this program to be accessible by anyone, anywhere, at anytime I want to develop something very similar to the webCAT to be integrated onto mobile platforms. I have considered two methods in this regard. First, I would like to develop another stand-alone program similar to the PC program I have already listed. Second, I want to develop a program that will overlay or interact with already existing mobile programs that the user already has.

These future steps, directions, and uses are only my first thoughts in relation to the webCAT. I am certain that during the process of development I will think of many more, some good and some bad. I believe that the webCAT can be an important tool in the advancement of creativity assessment and creative development.

Discussion

Additional Benefits

I have spent considerable time outlining practical uses of the webCAT in relation to creative products and in development of creativity within the individual. I would now like to talk about some other alternative uses and benefits I have envisioned for the webCAT.

First amongst these are benefits related to creative meta-cognition. Meta-cognition is the ability of the individual to self-regulate their cognitive processes in relation to a specific or general domain. In creating the webCAT I have paid particular attention to trying to assess creativity. I hope that when users work through the program there is an opportunity for understanding and growth in their ability to

self-regulate their creativity. Specifically, I hope users can establish a correlative link between their perceived level of expertise and self-assessment of their creative product in relation to expert assessment of the same creative product. This should allow users to identify and improve their strengths in a specific act or domain. There are many additional benefits to developing meta-cognition outside of creative self-regulation, many of which are related to an individual's self-determination and an individual's self-perception.

Creativity is a much talked about in education, however, what is not talked about enough is the relationship between individual creativity and self-determination; an individual's efficacy and motivation. Both efficacy and motivation are thought to be very important in individual development but all too often people forget that there is a strong interaction between an individual's efficacy and motivation and growth in a specific domain. I believe that the program I am creating will allow for individuals to be able to self-monitor and adjust their own efficacy and motivation dependent on the creative domain for which the webCAT is applied in. Allowing individuals to monitor ratings on their own creative works and more importantly to be active in interacting with raters and other creators is likely to have a positive influence on efficacy, especially if they can see improvements in creative performance. In regards to motivation, by creating an established interest in the domain through interaction it is much more likely that individuals will develop intrinsic motivation, which is likely to sustain growth both in a specific domain and in relative domains to the creative task. Even if there is not improvement in a specific domain an individual is much more likely to be able to identify strengths be able to apply these to other aspects of their life. And, in this action, individuals are much more likely to develop a positive self-perception in a creative domain.

There are many positive benefits for an individual who creates a positive self-perception. By using the webCAT to assess creativity I believe two of these perception related by-products are likely to

increase. They are (1) development of a positive and creative ideal-self and (2) an increase in self-esteem. When I refer to the development of the ideal-self I am referring to an individual's perception and understanding of, first, their ability level in a specific domain and, second, a realistic approximation of what they should be able to achieve in the same domain. With a positive ideal-self individuals will be more realistic in their approach to acts within a specific domain and not be overly influenced by small negative setbacks in their development. Self-esteem is fairly self-explanatory. In this instance it is important because self-esteem and specifically positive gains in self-esteem are essential to sustained development of creative potential. With a more positive self-perception of their creative development users should experience some improvement of self-esteem over time and thus be more likely to engage in future creative acts.

Questions and Concerns

Even with my grandiose vision of and extensive planning to implement this web-based program there are still many important and lingering questions to consider. One such questions is, "will the proposed model actually assess creativity?" Given the abundance of aggregate ratings websites and programs it is a very important question in determining the necessity for yet another rating program. To begin looking at this question I think it is important to first look at the proposed implementation of the Consensual Assessment Technique (CAT). There is no doubt that the CAT is a valid and reliable measure of creativity but will the proposed integration of the CAT just serve as an extension or would the validity and reliability need to be re-established? Personally, I believe in the later. New implementation of a proven system within a new domain will likely create conflicts and design changes which may or may not seriously change fundamental interactions in the system, i.e., using something old in a new way will likely change how the old thing works. I think it will be important to test reliability

and validity of the new implementation of the CAT with the old implementation of the CAT to make sure it works as it should in assessing creativity.

Another important question to consider is, “is the CAT the best method in a virtual environment to assess complex creative products?” Certainly there are many alternative methods to assessing creativity. Personally, however, I believe that the CAT is the most versatile method for assessing creativity in different products—I want to stress that it is the most versatile for assessing creative products, not general creative potential—and as such would be the best at this stage for implementation. However, depending on what aspect of creativity is to be assessed or if a specific community of practice would like a more robust assessment tool designed to work in an extremely limited capacity, it may be better to explore more specialized options. Also, it may be that in testing I will find that another assessment method offers greater flexibility; it is too early to tell but it is something worth paying attention to.

A final question to consider in implementation and with relation to the necessary components of the CAT to make it a successful creativity assessment tool is, “how do we attract raters to participate in rating creative products?” This is a very important question—maybe the most important—as the CAT requires raters of similar expertise – in a specific domain– to rate a creative product independently of each other and referenced with other creative products in that domain. By situating the webCAT into a specific community of practice I hope that raters will be more interested in furthering their community and would use that to serve as a motivator. I have also had suggestions of creating a 1:1 or similar barter system to insure that creators can find raters. Along these lines you can introduce an element of gamification to motivate. The real question then is, “what motivator will work best in different situations?” This is something I will seriously look into when attempting to integrate this program into different situations and communities. These lingering questions are important to consider while working

through the programming and implementation procedure and I will certainly keep them in mind as I work.

Conclusion

I firmly believe that my proposed webCAT will be an important step in the development of an accessible, easy-to-use, and adaptable creativity assessment tool. The benefits both explicitly stated and those that are implicitly implied are too great and important for an attempt not to be made. It is imperative to move creativity assessment into a digital and virtual space to make it both accessible and affordable for a new generation of creative individuals. A way must be made to insure that creativity assessment and creative development is brought to the forefront of psychological and development research. My proposed design is just another step in this process and I sincerely hope that whatever contributions I make will be thought-provoking and useful to researchers everywhere.

References

- Kaufman, J. D. (2010). Creativity stereotypes and the consensual assessment technique. *Creativity Research Journal*, 22(2), 200-205.
- Kaufman, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal*, 20(2), 171-178.
- Lebuda, I. M. (2013). Tell me your name and I'll tell you how creative your work is: Author's name and gender as factors influencing assessment of products' creativity in four different domains. *Creativity Research Journal*, 25(1), 137-142.
- Pretz, J. E., & Link, J. A. (2008). The creative task creator: A tool for the generation of customized, web-based creativity tasks. *Behavior Research Methods*, 40, 1129-1133.

Richard W. H. (2015). Feasibility of online divergent thinking assessment. *Computers in Human Behavior*, 46, 85-93. doi:10.1016/j.chb.2014.12.056