Verbal protocol analysis as a tool to understand the creative process

Paul T. Sowden, Department of Psychology, University of Winchester, UK

(paul.sowden@winchester.ac.uk)

Andrew Pringle, School of Psychology, Trinity College Dublin, Ireland (pringlea@tcd.ie)

Matthew Peacock, School of Psychology, University of Surrey, UK (m.peacock@surrey.ac.uk)

In press, *Handbook of Research Methods on Creativity*. Marc Stierand and Viktor Dörfler (Ed's).

Introduction

How do you gain insights into the mental processing of someone as they create?

This is one of the most intractable problems in creativity research. Whilst we can measure what goes into the creative process by considering the nature of the task and the individual difference characteristics of the creator, and we can observe its outputs in terms of the tangible products or ideas that are generated, what happens in between input and output has involved a great deal of introspection and inference. From this inference and introspection various ideas about the nature of the creative process have developed. One of the earliest formalisations was the descriptive, stage based account of Wallas (1926), which remains influential. Even now, much creativity research continues to focus on specific stages of the creative process such as the experience of insight that often accompanies creative problem solving (Kounios & Jung-Beeman, 2009) or distinct but mutually dependent stages of generating a number of creative ideas through divergent thinking and then evaluating or selecting them (Campbell, 1960; Ellamil, Dobson, Beeman & Kristoff, 2012; Guilford, 1950, 1956).

In our own work, we became interested in exploring the creative process as it actually unfolded so that we might examine the ways in which the sub-processes of generating ideas and evaluating them interact in order to produce a creative product (Cropley & Kaufman, 2012; Howard-Jones, 2002; Gabora & Ranjan, 2013). Neuroimaging approaches have made some headway at understanding this interaction in a lab-based context (e.g. Beaty, Benedek, Silvia & Schacter, 2016; Hao, Ku, Liu, Hu, Bodner, Grabner & Fink, 2016) but we were interested in how these processes unfolded outside the lab in the real world as people engaged with complex creative tasks over an extended period of time.

A different approach was therefore needed, one which could be used outside of the laboratory to directly observe the creative thinking process as it unfolded in all of its richness. One such approach that has been found to be generally useful for this kind of situation is the analysis of verbal protocols generated as participants 'think aloud'.

What is verbal protocol analysis?

The generation of verbal protocols requires participants to articulate their thinking process as they perform a given task (Ericsson & Simon, 1993). The protocols are recorded, transcribed, segmented into meaningful units and these are then coded into pre-defined categories that represent different task stages, content or processes. This allows statistical and other types of quantitative analysis and modelling to be applied, as well as qualitative analysis, to help researchers explore and understand what participants are actually doing as they engage with the task, and the findings can be related to the quality of the final product. Previous studies have used think-aloud methods to examine cognitive processes (Fayena-Tawil, Kozbelt & Sitaras, 2011) and activities which can be linked to cognitive processes (Atman, Chimka, Bursic & Nachtmann et al. 1999). Findings from this body of work suggest that verbal protocol analysis may indeed be a sufficiently sensitive technique to reveal the cognitive processes of participants performing creative tasks (Suwa & Tversky, 1997; Atman et al., 1999; Cross, Christiaans & Dorst, 1994 as cited in Cross, 2011).

For example, Atman and Bursic (1998) used a think aloud protocol to examine the performance of senior and first year engineering design students tasked with designing a playground. The verbal protocols were coded for different *design steps* such as 'problem definition or 'gathering information', the *information processed* such as safety guidelines or the age of the children, the *activity* such as reading or performing calculations, and the *object* or content such as playground layout or swings. Analysis provided a rich description of the students' approach to the task, for example revealing that they spend relatively little time gathering information and that senior students evidenced a higher frequency of transitions between different design steps than first year students. The frequency of transitions for both groups was also positively correlated with a measure of the quality of the final design (Atman et al. 1999).

Gilhooly, Fioratou, Anthony and Wynn (2007) used a think aloud design to study the processes underlying divergent thinking as measured using the Alternate Uses Task (AUT). Participants' responses were segmented into short phrases and these were then categorised. The initial coding scheme used 15 different codes to categorise the items such as 'episodic memory use' (where a possible use is stated with reference to a specific memory such as remembering one's grandfather using matchsticks as a wall-plug to secure a screw in a pre-drilled hole) or 'disassembly use' (such as removing the spokes from a bike wheel to use as knitting needles). Codes were then combined into four overarching strategies (retrieval from memory, using object properties to cue uses, conceptualising broad object use e.g. aesthetic, disassembly uses) and two processes (repeating object name to cue relevant info, repeating dominant or already given uses). By exploring the relationship between these strategies and processes and the fluency and novelty of the uses produced, Gilhooly et al. (2007) were able to identify that whereas fluency of production was more related to use of memory to retrieve already known uses, novelty of production was more related to a strategy of disassembly as well as to a lesser extent the use of memory. This example highlights the potential of a think-aloud method to drill down further than previous work into the sub-processes underlying different aspects of performance at alternate use generation and, more broadly, demonstrates the value of this approach for understanding the creative thinking process.

Verbal protocol analysis clearly has the potential to become a highly flexible tool for tackling a wide range of questions, addressing topics that range from a basic description of the creative process through to detailed understandings of the thinking processes involved. Examples might include:

- How long do people spend in different stages of the creative process and how does this vary with expertise?
- Do people move linearly through the different stages of the creative process or do they cycle back and forth, returning to previous stages, and how does this relate to the creativity of the final outcome?
- Do participants form an initial design and work on it or do they generate many before selecting one to focus on?
- Is spending more time in particular stages of the process (e.g. problem construction) predictive of the quality of the final product?
- How does the creative process change during an intervention to enhance creativity?

In the next section we run through more specifically the stages of conducting and analysing a verbal protocol study before moving on to work through a detailed example from our own research.

Steps to conduct a verbal protocol study

A useful initial step is to perform an analysis of the task to identify expected task components and processes. This provides a framework for the analysis of the subsequent think aloud data. For instance, you might be interested in the content that participants are working on, their cognitive processes, their affective processing, and the stage of the task or process. Each of these can be thought of as an overarching variable on which participants' responses can be coded. To achieve this, you would devise a coding framework that includes specific attributes that can be used to code for each variable. So for the content variable you might have attribute codes corresponding to the different content categories of interest. For the cognitive processes variable, you would code the different processes you wish to explore (e.g. retrieval of information from memory), for affective processing the expressed emotions that are of interest (e.g. you could code for expressed emotion present or absent, emotion valence, or even specific emotions) and for task stage the specific stages you wish to explore could be coded. Whilst developing your coding framework it is, of course, crucial to bear in mind the hypotheses that you intend to test and to ensure that your codes will enable you to test these.

With your pre-defined coding framework devised, you are ready to get going. Before you start the task proper it is wise to give your participants some practice at thinking aloud. For instance, you might get them to think aloud as they describe how many windows they have at their house or to plan a route somewhere. If the participant falls silent they can be prompted to "keep talking". Once you are confident that your participant is sufficiently comfortable and practised at thinking aloud whilst they conduct a task you can move on to the main task.

To keep track of your participant's thought process use audio and/or video recording. The latter can be very useful to capture tasks with visual elements such as graphic design. For example, revising, erasing or reworking visual content can help identify evaluative processes (Kozbelt, 2008). Further, if directed at the participant, the video recording can also be used to help fill in parts of the audio transcript that are unclear.

Once the task is completed the next phase is to transcribe the audio/video recordings. Unless your protocols are very short or you are an audio typist, a lot of time can be saved by employing a professional typist to undertake the transcription process, although if you do undertake the transcription yourself you may find this provides you with a useful first opportunity to engage with your data. There are computer programs available that may help you to speed up the transcription process by dictating the audio transcript that you are listening to into computer text.

Following transcription, the next step is to segment the transcripts into meaningful units. The granularity of the segmentation required will depend on your hypotheses and can be at the level of words, phrases, sentences, or even whole paragraphs. Segmentation is best carried out independently by two or more analysts with the results checked for agreement and differences reconciled through discussion.

The resultant segments are then coded according to your pre-defined coding scheme. Segments can be coded multiple times, once for each of your overarching variables. For instance, they might be assigned a code for content and another for thinking process. Once again the coding, or at least a sample of the coding (Gilhooly et al., 2007), should be conducted by at least two independent analysts and inter-rater reliability checked, with differences discussed and resolved as described above. It can be helpful to initially have the analysts code only one or two transcripts each, check for differences and resolve these before moving on to code the whole sample. This can help to uncover any differences in conception or approach to the coding and allow a consistent approach to be developed from the outset.

If you intend to explore the relationship between how participants approach the task and the quality of the ideas or products that they produce, then it is also necessary to score the final products. This can be done using any of the standard approaches to scoring the outcomes of a creative process such as the consensual assessment technique (Amabile, 1982) or scoring on dimensions such as uniqueness, fluency, flexibility and elaboration (Plucker, Qian & Wang, 2011; Silvia et al., 2008).

With coding of the data complete, and the final products scored, you will be ready to represent and visualise the data. There are numerous helpful ways to do this. One approach is to generate individual participant timelines of the task for each variable and to mark along the timeline blocks that represent continuous segments that have been assigned the same code (e.g. Atman & Bursic , 1998; Pringle & Sowden, 2017[a]). Thus, an impression can be gained of the time allocated by the participant to a given thinking process or to a particular stage of the creative process. Further, patterns of transition between these different processes or stages can be observed and analysed using methods such as Markov Chain analysis (Kaplan, 2008; Güss, 2018; see next section).

Another useful visualisation approach is to produce summary tables for each variable describing the percentage of time spent on each code. For example, calculating percentage of time spent in problem construction, idea generation and idea evaluation. This can give a broad brush impression of how participants approached the task. Finer granularity can be achieved by considering the conjunction of two or more variables and this can be usefully visualised through two or 3D plots (see Atman et al., 1999). For instance, the time spent engaging in a particular thinking process (associating ideas, retrieving information) as a function of the stage of the creative task. Conventional statistical tests of difference, estimation analyses, Bayesian or other approaches can then be used to test hypotheses regarding differences between groups, for instance, as a function of expertise or comparing before and after differences in an intervention study.

We next work through a particular implementation of a verbal protocol analysis in more detail to test specific hypotheses regarding the creative process.

An example study

At its core the creative process involves dual processes of associating information to form new ideas and analysing the results (Sowden, Pringle & Gabora, 2015). Our previous work, and that of others, suggested that individuals vary in the extent to which they are able to engage these associative and analytic processes in combination as opposed to serially and in their ability to shift between these processes (Beaty, Silvia, Nusbaum, Jauk & Benedek, 2014; Howard-Jones, 2002; Pringle & Sowden, 2017[b]). We were interested to explore the extent to which this shifting ability was important for the creative outcome in a realistic, complex creative task. We chose to work in the domain of garden design as this is an understudied yet rich creative domain that requires significant expertise and with which we were familiar. Full details of the study and the findings are reported elsewhere (Pringle & Sowden, 2017[a]). Here we focus on the basic rationale and methods as a detailed example of the application of verbal protocol analysis to test hypotheses about the creative thinking process. Specifically, we hypothesised that frequency of shifting between associative and analytic processes would predict the quality and creativity of the final garden designs. Further, following on from the work of Dietrich (2004), we also hypothesised that affective processing would be related to the creativity of the final product.

To test these hypotheses, we needed to be able to track participants' thinking processes as they happened, a task for which the generation of verbal protocols for analysis through participants' thinking aloud is well suited. To explore whether shifting ability varied with expertise, we compared professional garden designers with students. We also included a 'creative' control group of fine artists with highly developed drawing skills like the designers, but no specific knowledge of garden design, and a group of professionals with no particular drawing or garden design expertise.

Working with a professional garden designer, we devised a brief that asked our participants to produce a design for a garden on A3 paper within a period of forty-five minutes. Specifically, participants were asked to produce a design for a garden "*based on a journey and the series of experiences those who walk around the garden will have on this journey*". The brief asked them to make the garden as creative as they could but that it should also be appropriate and work in the context of the brief. Participants were also briefed on factors such as the site dimensions, aspect (that is the compass direction the plot faces) and available budget and provided with multiple sheets of paper to draft on before producing their final design. Full details of the brief can be seen in Pringle (2015) and Pringle & Sowden (2017[a]).

Participants were asked to 'think aloud' as they worked on the garden design task and given the following instructions to help them to understand how they should do this:

"*While designing you will be asked to 'think aloud' your thoughts, which means you will be asked to speak out whatever you are thinking at the time. I would like you to focus on describing what you are thinking. Don't worry about complete sentences and don't hold back from describing hunches, guesses, wild ideas, images, plans or goals that you have. Don't over explain or justify. Analyse no more than you would normally. Just describe whatever is on your mind at the time. Try and get into the pattern of saying what you're thinking about now, not of thinking for a while and then describing your thoughts. Though I am present you are not talking to me. Instead you are to perform this task as if you are talking aloud to yourself. Speak as continuously as possible, try to say something at least once every 15 seconds, even if only "I'm drawing a blank". Try and also speak audibly, watch out for your voice dropping as you become involved. I will prompt you as we go along to help you to think aloud with phrases such as "what are you thinking now?" and "can you speak up?" if your voice drops*".

They were then given two practice tasks to get them used to thinking aloud. These were to think-aloud while they answered the question “what is the sixth letter after B?” and to ‘think aloud’ while naming ten animals.

The process of designing the garden was recorded using a digital high-definition Sony video camera. Transana software (Woods & Fassnacht, 2012) was used to analyse the audio and video data captured by the camera. This package enabled segments in the video to be linked to segments in the participants’ verbal reports so that both the video and audio data could be used to facilitate coding for attributes of the different modes of thinking within the verbal protocol.

The verbal protocols were transcribed and segmented. One segment was defined as a sequence of words, phrases or sentences of any length that made up one distinct statement about something, such as an idea or topic (Suwa & Tversky, 1997; Atman et al.,1999; Gilhooly et al., 2007). Typically, segments were 5 to 10 seconds long. A total of 13,611 segments were coded across the 48 participants.

Following segmentation, the segments were coded independently by two coders. The experimenter coded all segments across all participants using a pre-defined coding scheme (Table 1). The coding scheme was based on dual-process models of creativity (Gabora, & Ranjan, 2013; Howard-Jones, 2002) and dual-process models of cognition (Epstein, 2003; Evans & Stanovich, 2013; Frankish, 2010; Kaufman, 2011).

Table 1. Examples of variables and attribute codes used to code segments within the verbal protocol. The full coding scheme can be seen in Pringle & Sowden (2017[a]).

Overarching variable	Coding attribute	Definition of attribute	Example of attribute in protocol
Associative mode of thinking	Memory retrieval	Making associations to knowledge and/or prior experiences (but not evaluating it)	*e.g. This reminds me of the landscape architect George Hargreaves*
	Generating ideas/concepts	Any new ideas or elements of new ideas produced	*e.g. What about a stream here, a water feature there*
	Insight moment	Moment of sudden knowing	*e.g. Aha I know what I can do here*
Analytic mode of thinking	Logical deduction	Deduction of causal relationships between elements	*e.g. The scale is x metres so this feature will have to be y metres*
	Evaluating remembered experiences/past behaviour	Evaluating remembered information about past design relevant experiences	*e.g. That decision in the past was going against my grain*
	Evaluation of design ideas/concepts	Evaluating ideas, evaluating in the context of something else (e.g. brief, expectations).	*e.g. That's working/that's not working. That's not going to work within the scale.*
Two modes meshed together	Two modes meshed together (associative cognitive & analytic cognitive)	Segment clearly has both associative cognitive and analytic cognitive modes within it	*e.g. I can't have a mountain (no previous mention of mountain)*
	Two modes meshed together (associative cognitive & analytic affective)	Segment clearly has both associative cognitive and analytic affective modes within it	*e.g. It's always more interesting if there is a rise in the garden, level sites are a bit boring*
Unknown	Documentation	A statement without analytical/evaluative or generative comment	*e.g. So it says in the brief to make the design as creative as possible*
	Reminder to speak	Participant is asked to 'keep talking' or to 'speak up'	

Attributes that reflected the operation of associative and analytic modes of thinking were translated into examples of activities which could then be coded for in the verbal protocol as associative or analytic. Only attributes that distinguished between the two modes of thinking were included. Segments were coded based on the attributes they contained and the mode of thinking they reflected. A segment could be coded for multiple attributes from the same over-arching mode. For example, a segment could be coded with “generating ideas/concepts” and “images, metaphors and analogies” from the overarching “associative mode”. The initial aim was to also code each segment as only reflecting the operation of one overarching mode of thinking; associative or analytic. However, it was not always possible to code segments as predominantly containing associative or analytic content, sometimes segments appeared to reflect similar levels of both modes of thinking. A code was therefore included in the coding scheme to reflect the operation of different modes of thinking within the same segment, labelled as two modes meshed together. For example, in the segment “it’s not going to be curved because that doesn’t work” the participant introduces and evaluates the idea of a curve but it is not clear whether the idea or evaluation came first. There were two different types of two modes meshed together segments; one which involved the operation of both associative cognitive and analytic cognitive modes and the other which involved the operation of both associative cognitive and analytic affective modes.

Some segments contained no content, or a lack of clear content. These segments were labelled as either documentation, reminder to speak or experimenter talk as appropriate. For example, in the segment “*so it says in the brief to make the design as creative as possible*” there is no reference to any attributes of either associative or analytic thinking, the participant is simply re-reading the brief they were given.

Finally, a range of evidence suggests that evaluative processing that draws on analytical thinking can be cognitive or affective (Dietrich, 2004). Consequently, drawing on Dietrich's (2004) framework, we coded for the operation of associative and analytic processing on two types of content; cognitive and affective (Table 2). This resulted in four categories of thinking: associative-cognitive, associative-affective, analytic-cognitive and analytic-affective. Words in the associative-affective, analytic-affective and two-modes meshed together associative-cognitive/analytic-affective segments, appearing to reflect affective content, were identified. These words were checked against Warriner, Kuperman & Brysbaert's (2013) database of norms for the affective meaning of words in order to check the validity of the coder's subjective judgment. The database contains 13,915 words that have been rated by individuals on a scale of 1 to 9 on dimensions of valence, arousal and dominance. Only words rated at least one standard deviation above or below the mean on at least one dimension of affect retained the affective code.

Table 2. Showing the four modes of thinking coded for within verbal protocols in the present study and their alignment with Dietrich's (2004) four modes of thinking

Mode of thinking proposed	**Corresponding mode of thinking in Dietrich's (2004) framework**
Associative-cognitive	Spontaneous cognitive
Associative-affective	Spontaneous emotional
Analytic-cognitive	Deliberate cognitive
Analytic-affective	Deliberate emotional

Coding was checked for reliability by an expert second coder. Random segments to be second-coded were chosen by first dividing the protocol of each participant into four distinct time periods; 0-12, 12-24, 24-36 and 36-48 minutes. A chunk of consecutive segments from one of these four time periods of a participant's verbal protocol was then chosen for the second coder to code. Disagreements between coders on the coding categories to apply were discussed. After coding of all segments was completed, the first coder checked through the coding across all segments to make sure the codes were applied consistently and any disagreements between coders were resolved.

Following completion of the study, three judges with experience of judging at garden design shows in the United Kingdom, including the Chelsea Flower Show, were recruited to rate the designs using Amabile's (1982) consensual assessment technique (CAT). Judges were presented with original copies of all sketches of all designs produced by all participants. Judges rated designs blind to which groups produced which designs. They were asked to rate designs based on three dimensions; brief, design and creativity/wow factor. "Brief" referred to how well the designs met the requirements of the brief while "design" referred to the quality of the design that was evident in design sketches. The "creativity/wow factor" was the creativity that judges saw evident in the designs. Judges were to assess each dimension as separately as possible. Designs were rated relative to one another on each dimension rather than against some absolute standard for garden design. Judges rated designs using a 1 to 5 point scale. Each judge rated designs in a random order, allocated by the experimenter and were instructed to make full use of the 1 to 5 point scale.

In order to test our hypotheses about the extent to which shifting between modes of thinking predicted creative outcomes, we needed a method of analysis that allowed us to explore the shifts between the different types of thinking process (states) that occurred in order to test which types of shift were predictive of creative outcomes. Assessing transitions between different states is a problem well suited to Markov Chain analysis.

A Markov chain is a model of a sequence of categorical events that evolves over time (Kaplan, 2008). The model assumes that the sequence is stochastic and that the probability of the current categorical event depends only on the categorical event immediately prior to it (Kaplan, 2008). Thus, given a particular event type has occurred, one can calculate the probability that each of a range of other event types will follow it and compare the relative probabilities. This provides insights into the likely sequence of events and allows patterns of event transition probabilities to be related to final creative outcomes to see if there is a relationship between creativity and the frequency of specific transition types.

So, taking the example of the associative and analytical categories coded for in the present study, if events were randomly distributed then there is a .5 probability that the current mode is associative and a .5 probability that it is analytic. There is a .5 probability that the mode immediately following the current mode is associative and a .5 probability that the mode immediately following the current mode is analytic. Consequently, there are four possible types of transition, associative to associative, analytic to analytic, associative to analytic and analytic to associative. Maintaining the assumption of randomness, the probability of each of these types of transition occurring is .25. Transition probabilities thus sum to 1 (Kaplan, 2008). However, in reality we might expect that events would not be randomly distributed and that transition probabilities would vary for individuals and between groups. Thus, to calculate the actual probabilities, generalising for the example of associative to analytic transitions, we can say:

Transition probability (associative to analytic) = Σ (associative to analytic) / Σ (associative to analytic + associative to associative)

Conventional statistical analyses such as Analyses of Variance can then be used to compare different types of transition probabilities across groups and regression can be used to predict creative outcomes as a function of transition probabilities. As an example, in our study we calculated the probability of transitions between the four different categories derived from Dietrich's (2004) model (Table 2).

Analysis of variance revealed that professional and student garden designers showed a greater tendency to shift from analytic-affective to associative-cognitive compared to controls, suggesting that for the designers an affective reaction to some aspect of their design was more likely to trigger a further process of association making to generate additional ideas. Intriguingly, a similar finding on the benefits of idea evaluation for subsequent idea generation has been reported in a lab based neuroimaging study (Hao, Ku, Liu, Hu, Bodner, Grabner & Fink, 2016). This illustrates how the think-aloud method could be a useful means of determining the ecological validity of ideas about creative thinking garnered from neuroimaging work or as a means of generating novel hypotheses to test in neuroscientific studies of the creative process.

Limitations

Although verbal protocol analysis is a very valuable approach to exploring creative processes, as with all methods it also has some drawbacks and limitations. Perhaps foremost amongst these is simply that it is a very time consuming research approach. The time required for the processes of transcription, segmentation, coding and inter-rater agreement, necessary to generate quantitative data for analysis, is substantial. Consider, however, that it may not always be necessary to analyse a full transcript in order to do interesting work. For instance, it may be that you are only interested in a specific time window, such as problem construction at the start of the creative process. Alternatively, it may be possible to gain useful information by analysing randomly selected time windows extracted from the overall time course.

Other issues, often raised, are whether the process of thinking aloud may actually change the thinking process and also whether it even provides a true reflection of the underlying thought processes. For example, it has been suggested that the requirement to verbally report on internal processes may interfere with concurrent task performance (e.g. Schooler, Ohlsson & Brooks, 1993). This has been termed verbal overshadowing. However, other work has suggested that thinking aloud does not interfere with task performance (e.g. Gilhooly et al., 2007) and in an extensive review Ericsson and Simon (1993) argue that the sequence of participants' thoughts is unaffected by thinking aloud, although the total time on task can be increased by the requirement to think aloud during task performance (Fox, Ericsson & Best, 2011). The key point is that to ensure thinking aloud does not change task performance, participants are merely asked to verbalise their thoughts, not to interpret or reflect on their thinking processes. When participants are required to engage in such explanation the evidence suggests this does alter task performance (Ericsson & Simon, 1993). Thus, it is crucial that participants are given the opportunity to practise thinking aloud without further interpretation required.

If you are concerned about whether thinking aloud might interfere with your task Gilhooly et al (2007) suggest the use of a control group who perform the same task but without thinking aloud. For example, in their work, one group of participants thought aloud as they generated uses for six different objects. A control group performed the same task without thinking aloud. The number of uses generated and their novelty did not differ between groups, suggesting that having to think aloud does not overshadow performance on alternate use generation tasks.

A final limitation worth stating is that think-aloud protocols are evidently only capable of revealing conscious reasoning and not unconscious processes (Allen & Thomas, 2011).

Summary

Analysis of verbal protocols generated by asking participants to think aloud can be a very valuable tool for understanding the creative process. It has the advantage of being usable in almost any creative situation, including extended real-world creative tasks, and thus is adaptable to non-laboratory settings. Evidence suggests that so long as participants are simply required to verbalise their thoughts without further interpretation then thinking aloud provides a valid and reliable approach for revealing conscious sequences of thinking, which can then be related to creative task performance and outcomes.

References

Allen, A. P., & Thomas, K. E. (2011). A dual process account of creative thinking. *Creativity Research Journal, 23*(2), 109-118.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*, 997-1013.

Atman, C. J. and Bursic, K. M. (1998). Verbal Protocol Analysis as a Method to Document Engineering Student Design Processes. *Journal of Engineering Education*, *87*, 121-132.

Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. *Design Studies, 20*(2), 131-152.

Beaty, R.E., Benedek, M., Silvia, P.J., & Schacter, D.L. (2016). Creative Cognition and Brain Network Dynamics. *Trends in Cognitive Sciences, 20,* 87-95.

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42*, 1186 –1197.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. Psychological Review, 67, 380-400.

Cropley, D.H., & Kaufman, J.C. (2012). Measuring functional creativity: Non-expert raters and the creative solution diagnosis scale. *The Journal of Creative Behavior, 46*(2), 119-137.

Cross, N. (2011). *Design thinking: Understanding how designers think and work.* NY, USA: Berg.

Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic Bulletin & Review, 11*, 1011–1026.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis*. Cambridge, MA: MIT Press.

Ellamil, M., Dobson, C., Beeman, M., & Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage*, *59*(2), 1783-1794.

Epstein, S. (2003). Cognitive-experiential self-theory of personality. In I.B. Wiener, T. Millon, & M. J. Lerner (Eds.), *Handbook of psychology: Personality and social psychology* (pp. 159– 182). Hoboken, NJ: John Wiley & sons Inc.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223–241.

Fayena-Tawil, F., Kozbelt, A., & Sitaras, L. (2011). Think global, act local: A protocol analysis comparison of artists' and non-artists' cognitions, metacognitions, and evaluations while drawing. *Psychology of Aesthetics, Creativity, and the Arts*, *5*, 135–145.

Frankish, K. (2010). Dual-process and dual-system theories of reasoning. *Philosophy Compass*, *5*, 914 –926.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137*(2), 316-344.

Gabora, L. & Ranjan, A. (2013). How insight emerges in a distributed, content addressable memory. In A. Bristol, O. Vartanian & J. C. Kaufman (Eds.), *The Neuroscience of Creativity* (pp. 19-44). NY: Oxford University Press.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. British Journal of Psychology, *98*, 611–625.

Guilford, J. P. (1950). Creativity. American Psychologist, *5*, 444-454.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*, 267-293.

Güss, C.D. (2018). What is going through your mind? Thinking aloud as a method in cross-cultural psychology. *Frontiers of Psychology, 9.*

Hao, N., Ku,Y., Liu, M., Hu, Y., Bodner, M., Grabner, R.H. & Fink (2016). Reflection enhances creativity: Beneficial effects of idea evaluation on idea generation. *Brain and Cognition, 103,* 30-37.

Howard-Jones, P. A. (2002). A dual-state model of creative cognition for supporting strategies that foster creativity in the classroom. *International Journal of Technology and Design Education, 12*, 215–226.

Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. *Developmental Psychology*, *44*, 457–467.

Kaufman, S. B. (2011). Intelligence and the cognitive unconscious. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 442–467). New York, NY: Cambridge University Press.

Kounios, J., & Jung-Beeman, M. (2009). Aha! The cognitive neuroscience of insight. *Current Directions in Psychological Science, 18,* 210-216.

Kozbelt, A. (2008). Hierarchical linear modeling of creative artists' problem solving behaviors. *Journal of Creative Behavior*, *42*, 181–200.

Plucker, J. A., Qian, M., & Wang, S. (2011). Is Originality in the Eye of the Beholder? Comparison of Scoring Techniques in the Assessment of Divergent Thinking. *Journal of Creative Behavior*, *45*, 1-22.

Pringle, A. (2015). Creative thinking: a mode shifting hypothesis. *PhD Thesis.* University of Surrey, Guildford, UK.

Pringle, A., & Sowden, P. T. (2017)[a]. Unearthing the creative thinking process: Fresh insights from a think-aloud study of garden design. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 344–358.

Pringle, A., & Sowden, P. T. (2017)[b]. The Mode Shifting Index (MSI): A new measure of the creative thinking skill of shifting between associative and analytic thinking. *Thinking Skills and Creativity, 23*, 17–28.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*, 166-183

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*¸ 68-85.

Sowden, P. T., Pringle, A., & Gabora, L. (2015). The shifting sands of creative thinking: Connections to dual process theory. *Thinking & Reasoning, 21*, 40–60.

Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. *Design Studies*, *18*, 385–403.

Wallas, G. (1926). *The Art of Thought*. Watts & Co.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*, 1191–1207.

Woods, D., & Fassnacht, C. (2012). Transana (Version 2.51) [Computer software]. Madison, WI: The Board of Regents of the University of Wisconsin System. Retrieved from http://www.transana.org