Improving assessment of analyst-relevant divergent thinking
Test validation, automated scoring, and brain signature

Henk J. Haarmann, PhD
Melody S. Berens, PhD
Polly O’Rourke, PhD
Sergey Blok, PhD
Alexei Smaliy, BA
Timothy G. George, MA
Kristin Grunewald, BS
Janet Cook, MA
Joseph Dien, PhD
Suzanne Freynik, MA

PURPOSE—To validate CASL’s Analyst-Relevant Divergent Thinking Task (ARDTT), determine the feasibility of automatically scoring parts of the ARDTT, and identify a brain signature of good verbal divergent thinking.

CONCLUSIONS—The ARDTT provides a valid divergent thinking measure, and automatically scoring parts of the ARDTT is feasible. Good verbal divergent thinking is indicated by increased active processing in the left compared to right hemisphere of the brain.

RELEVANCE—A valid ARDTT and automated scoring enable better assessment and evaluation of training programs for improving analysts’ divergent thinking. Such programs may be further enhanced by identifying a method for increasing verbal divergent thinking through greater left-hemisphere engagement.

Executive Summary

PURPOSE

Language analysts and other intelligence analysts must have good divergent thinking—that is, the ability to think of all relevant possibilities during problem solving. Recently, the University of Maryland Center for Advanced Study of Language (CASL) piloted and delivered an interactive one-day course that teaches analysts cognitive strategies to improve their divergent thinking. CASL also recently delivered the materials for an Analyst-Relevant Divergent Thinking Task (ARDTT). The broad aim of the present study was to address the following three questions related to the ARDTT and the divergent thinking course:

(A) Is the ARDTT a valid test that (1) consists of items that are rated as highly relevant to analysis, (2) correlates positively with an existing measure of general divergent thinking ability, and (3) can be divided into two highly correlated tests for future use in course evaluation? A positive answer to these questions would motivate the use of ARDTT in future studies aimed at providing evidence that the divergent thinking course or other methods improve analyst-relevant divergent thinking.

(B) Is it feasible to automatically score the flexibility of a person’s divergent thinking by computer? Flexibility is a key aspect of divergent thinking performance that reflects the number of categories or domains in which ideas are generated (e.g., use of an airplane as a transportation vehicle versus missile). Automated scoring is challenging since divergent thinking tests have no finite set of correct responses that can be known prior to testing. Nevertheless, automated scoring is worthwhile to develop since human scoring of responses is very time intensive.

(C) Is there a neurophysiological signature of good verbal divergent thinking, and can this signature be identified with a simple noninvasive method for changing brain waves (e.g., listening to periodically repeated tones or acoustic beats)? Methods for changing brain waves that are inspired by the neurophysiological results could be incorporated in a neural-inspired version of the divergent thinking course.

CONCLUSIONS

CASL’s research led to affirmative answers to all of the above three questions and subquestions:
The ARDTT is a valid test that analysts rate as highly relevant to their work, and the test can be divided into two parallel forms. Automated computer scoring of flexibility is feasible; however, the computer scoring methods must be optimized before implementation. Engaging the left hemisphere of the brain more than the right provides a neural signature for greater originality and fluency of divergent thinking. Alpha acoustic beats—that is, tones presented 10 times per second—are useful for interpreting the neural signature but may not improve verbal divergent thinking.

**RELEVANCE**

Because the ARDTT accurately measures divergent thinking and because analysts rated its items as highly relevant to their work, the ARDTT could replace tests of general divergent thinking in CASL’s divergent thinking course and assessments of analysts for other purposes.

With two parallel forms of the ARDTT, researchers can evaluate changes in verbal divergent thinking (e.g., after a training course) that cannot be explained by a trivial practice effect, that is, repeated practice with the same test items. Further, automated scoring of the ARDTT will reduce the number of hours human coders must spend scoring tests and will therefore make large-scale testing possible.

To help analysts who must find solutions under time pressure, future studies should evaluate simple methods for quickly engaging the left hemisphere of the brain. Such quick engagement may induce a pattern of brain activity associated with greater originality and fluency of divergent thinking.
Executive Report

This report discusses CASL's most recent research on divergent thinking, in three parts:

PART A  Developing and validating the Analyst-Relevant Divergent Thinking Task (ARDTT) (page iii)

PART B  Determining the feasibility of automatically scoring parts of the ARDTT (page iv)

PART C  Identifying the neural signature of good divergent thinking with a brain wave method (page vi)

PART A

Developing and validating the Analyst-Relevant Divergent Thinking Task

PURPOSE

Creative thinking is valuable and beneficial to the performance of many activities, including product invention, business strategies, and, importantly, analysis. To measure creative thinking skills, researchers have designed several types of tests, such as the Torrance Tests for Creative Thinking (TTCT). These existing creative thinking tests tap into general creative thinking abilities; however, these general creative thinking tests are not necessarily valid for any specific profession. In other words, they have low "face validity"—they may not measure what they intend to measure within a profession.

This low face validity makes it difficult to draw conclusions about the impact of creative thinking within a field. Thus, while researchers might be able to provide experimental manipulations or interventions that show an improvement on general creative thinking abilities (e.g., CASL's divergent thinking course), they may not be able to determine whether these same experimental manipulations would translate to improvements on creative thinking abilities in a particular field, such as analysis.

What this study investigated

This study focused on one type of creative thinking—divergent thinking—which is the ability to think of all relevant solutions during problem solving. To show that improved divergent thinking is beneficial specifically for government analysts' work performance, CASL designed a divergent thinking task that is relevant to the type of work analysts perform in the workplace: the Analyst-Relevant Divergent Thinking Task (ARDTT). In addition to ensuring that the ARDTT is relevant to analysts (i.e., has face validity), CASL researchers took steps to ensure that the ARDTT indeed measures the psychological process of divergent thinking.

CONCLUSIONS

CASL's divergent thinking task research supports the following conclusions:

1. Expert analysts judged the ARDTT to be highly relevant to their work.

   CASL first consulted with four government subject matter experts (SMEs) in language analysis to identify the crucial abilities and types of problems faced by professional analysts. The SMEs identified the following three constructs as being most relevant to their work: broad thought about meaning (i.e., interpreting coded communications), understanding long-term ramifications (i.e., identifying consequences), and scenario interpretation.

   Following this guidance, CASL drafted the ARDTT—modeling it after the verbal form of the TTCT, in which participants are presented with situations and instructed to ask questions, hypothesize about what might be happening or going to happen, and provide alternative responses. CASL then asked the SMEs to rate each question and the ARDTT overall on a 5-point scale judging the relevance to analysis work, where 1 = not at all relevant and 5 = highly relevant.

   Results showed that overall, SMEs rated each construct as 4 or higher, and that the task items received an average rating of 4.2, indicating that SMEs believed that the task and items on the task to be relevant to the work of analysis. Based on SME comments, CASL replaced or revised several of the questions and then created a computerized version of the test.

2. The ARDTT successfully measures divergent thinking.

   To ensure the ARDTT measures the psychological process of divergent thinking, CASL administered the ARDTT and the Abbreviated Torrance Test for Adults (ATTA) in the same session and then correlated the scores. The correlation between the overall scores of the ARDTT and ATTA showed a moderately significant relationship between these tasks, indicating that these tasks both measure divergent thinking.

   Correlations between the ARDTT and ATTA subtests showed that the ARDTT subtests have the strongest relationship with ATTA's subtest for verbal divergent thinking. This result is not surprising because all ARDTT subtests were modeled after the verbal form of the TTCT. These results indicate that the ARDTT provides a measure of the psychological process of verbal divergent thinking.

   CASL created two parallel forms of the ARDTT (Forms A and B), so they could be used, for example, in future research that compares a person's
ARDTT scores before an experimental intervention (such as divergent thinking training) to their scores after an intervention. Statistical analyses show that the two versions of the ARDTT are highly correlated and have good internal consistency (i.e., a form of test reliability).

RELEVANCE

The ARDTT could replace tests of general divergent thinking ability, such as the Torrance Test of Creative Thinking, in the evaluation of an analyst divergent thinking course and in assessments of analysts’ cognitive abilities for other purposes.

Further, because analysts rated the ARDTT as highly relevant to their work, it is more likely that individual differences in divergent thinking ability (e.g., as a function of job experience) and changes in divergent thinking ability (e.g., as a function of training), as assessed by the ARDTT, have an impact in the analyst workplace. Analysts will also likely be more motivated to take an analyst-relevant test, which in turn will provide a better assessment of their abilities.

The availability of two parallel forms of the ARDTT enables researchers to evaluate changes in verbal divergent thinking (e.g., due to a course) that cannot be explained by a practice effect—that is, repeated practice with the same test items.

CONSIDERATIONS

Establishing test norms: In this first study, our participants were undergraduate students, not government analysts. To apply the ARDTT in the analytic workplace, test norms should be developed for the population of analysts and for the general adult population. Such norms enable comparison of the score of an individual to the performance of the group of individuals who are representative of the population (e.g., the score of a particular test taker might be better than the score of 85 percent of the individuals in that group).

Measuring verbal divergent thinking: Given the modest correlation between the ARDTT and the ATTA, further research is needed to demonstrate that the ARDTT is a good measure of the psychological process of verbal divergent thinking. This evidence could come, for example, from studies that demonstrate that ARDTT performance is affected by factors that are known to have an impact on the causal mechanisms of divergent thinking, such as the role of executive control and self-regulatory motivation. The correlation between ARDTT and ATTA in this study may have been modest in size for several reasons, such as differences in test administration procedure (computer-based and paper-and-pencil, respectively) and fatigue that develops in a long testing session. The ARDTT provides a foundation for conducting a wide variety of studies aimed at evaluating and improving divergent thinking in government analysts.

Comparing the ARDTT to the Torrance test: A trade-off between the ARDTT and the (abbreviated and full version of) Torrance test of divergent thinking is that the ARDTT (but not Torrance test) has demonstrated analyst-relevant content, whereas the response forms of the Torrance Test (but not ARDTT) can be sent in to the Scholastic Testing Service for fee-based scoring by human coders, with a faster turnaround time (i.e., 3 to 4 weeks) than human scoring of the ARDTT. Automated scoring by computer, which hitherto has been unavailable, would save money and time.
of responses to questions that test a participant’s creativity.

Flexibility of divergent thinking is defined in terms of the number of semantic categories—that is, domains of thought (e.g., furniture, clothing, food, fuel)—in which a person generates ideas. Greater flexibility of divergent thinking is indicated by a greater number of semantic categories.

The primary goal was to develop models that evaluate the flexibility of test takers’ responses and produce scores that significantly correlate with human judgments. To be useful, the scoring method should not involve extensive human preprocessing and should have the potential to serve as an out-of-the-box solution. In addition, the method should apply to a variety of test items, including never-before-seen items, not only to a predefined set of test items.

Approach

The two models selected for this study are (1) the supervised-learning approach, in which the model attempts to learn the categories used by the human coders, and (2) an unsupervised learning approach, in which the model attempts to determine the diversity of responses for each participant in the semantic space of words he or she uses. A semantic space is a mathematical representation of a large body of text. For this approach, no prior human input is necessary and the model can be applied to any arbitrary creativity test item. The supervised approach tends to be more accurate, while the unsupervised approach tends to be more readily applied to new problems.

Each computer model scored responses to divergent thinking tests from two previous studies that had already been human-coded for fluency, flexibility, and originality.

CONCLUSIONS

1. Automatically scoring the flexibility of responses on divergent thinking tests is feasible.

The models’ predictions and the human-coded flexibility judgments correlated strongly. Overall, the supervised models performed more accurately and more consistently, while the unsupervised models, although showing promise, depend to a substantial extent on the particulars of their implementation, as described in the Technical Details section of this report.

2. The models must be optimized before implementation in actual testing situations.

The models and the data correlated significantly in this study; however, the correlations are too low for immediate practical implementation, and additional work geared toward optimizing model performance is needed. For example, future research needs to add the size of the training dataset. Although the training dataset in this study consisted of 500 to 700 responses per scenario (nearly all the available data), it is not known whether more data would make the models perform better or if and when diminishing returns are obtained.

RELEVANCE

This investigation shows that it is feasible to develop computer programs that automatically score the flexibility of a person’s divergent thinking. It also demonstrates that it is feasible to adopt a variety of computer modeling approaches for this effort, providing a fertile ground on which to base further development of the automated scoring of the flexibility of divergent thinking.

Automated scoring reduces the human effort and time and thus also the cost involved in determining test performance, which in turn enhances the prospect of providing the Intelligence Community with divergent thinking assessment and evidence-based training. In addition, automated scoring is likely to increase the benefits of divergent thinking training, since it enables faster learner feedback, which promotes better skill acquisition.

CONSIDERATIONS

Enhancing automated scoring: Automatically scoring divergent thinking could be further enhanced in two ways. First, automated scoring, which this investigation found to be feasible for the flexibility of divergent thinking, may also be feasible for determining the originality of divergent thinking. Originality of divergent thinking refers to the frequency with which an idea is generated among a larger group of individuals, such that an idea is considered more original if fewer individuals generate it. Originality is relevant to consider for analysts, since our adversaries often use tactics that involve ideas that we are least likely to think of (e.g., the use of a large commercial airplane as a powerful missile in the now familiar 9/11 scenario).

Second, automated scoring could be improved by including corpora with topics that are specifically chosen to be relevant for the scenarios that the test taker is thinking about. Topic relevancy is important to consider, because human problem solving is greatly influenced by the context in which it takes place, including the topics in one’s mind. These topics may also be feasible for determining test performance, which in turn enhances the prospect of providing the Intelligence Community with divergent thinking assessment and evidence-based training. In addition, automated scoring is likely to increase the benefits of divergent thinking training, since it enables faster learner feedback, which promotes better skill acquisition.

© 2011 University of Maryland. All rights reserved. December 2011
Supplying relevant associates for problem solving: The associative knowledge structures used in automated analyses of corpora have the potential to support the goal of improving analyst divergent thinking in two ways. (1) Associative knowledge structures can be used in automated scoring approaches as the present study demonstrated. (2) They might also be used as part of a method for directly improving an analyst’s divergent thinking on the job. Specifically, computer-derived associative knowledge structures may be used to provide an analyst with weak but relevant associates. These relevant associates could enable an analyst to better generate alternative solutions—not just those that come easily to mind. It would be worthwhile to empirically validate the usefulness of associative knowledge structures to analysts’ abilities to generate more possible solutions.

PART C
Identifying the neural signature of good divergent thinking with a brain wave method

PURPOSE
Because good divergent thinking is important to the work of language analysts and other intelligence analysts, the University of Maryland Center for Advanced Study of Language (CASL) recently piloted and delivered an interactive one-day course in which analysts learn to apply cognitive strategies to improve their divergent thinking.

Two findings from recent cognitive neuroscience research suggest that it may be feasible to enhance the course with a neural-inspired approach for improving divergent thinking.

The first finding concerns the neural signature for good divergent thinking. Results from a recent study\textsuperscript{10} revealed that the generation of highly original responses is associated with a decrease in the amount of active processing in the brain’s right hemisphere, as indicated by increased alpha brain waves. Alpha brain waves are brain waves that wax and wane about 10 times per second; the presence of these brain waves reflects that a person is in an alert but calm and receptive state.

The second finding is that there may be a relatively simple-to-apply method for temporarily and rapidly increasing a person’s alpha brain waves without any prior training. This method, which is called alpha acoustic beats, involves listening to tones that are presented 10 times per second for a few minutes immediately before or during a cognitive task.

What this study investigated
The major aims of this investigation were (1) to replicate and refine the description of the neural signature of good divergent thinking, (2) to determine whether alpha acoustic beats provide a useful method for identifying this signature, and (3) to explore whether alpha acoustic beats can be used for improving divergent thinking. The second and third aims were motivated by the association of good divergent thinking with increased alpha brain waves. A neural signature that can be identified and rapidly enhanced with a simple method, such as listening to alpha acoustic beats, could then be used in a training course aimed at improving analyst divergent thinking.

Approach
University of Maryland students participated in two experiments in which the electrical activity of their brains was recorded with electroencephalography (EEG).

Experiment 1: Participants were exposed to different types of acoustic stimulation while engaged or not engaged in a cognitive task (i.e., metaphor judgment). Figure 1 illustrates the different types of acoustic stimulation.

A neural signature that can be identified and rapidly enhanced with a simple method, such as listening to alpha acoustic beats, could then be used in a training course aimed at improving analyst divergent thinking.
metaphor judgment task had a positive impact on performance on that task. Beta acoustic beats were included for exploratory reasons only and did not generate any effect.

Experiment 2: Participants listened to alpha acoustic beats and pink noise, or to pink noise only, with and without a cognitive task. The cognitive task in this study involved generating responses to several items of a divergent thinking test. For example, participants viewed on a computer screen the test item “a light in darkness” (Figure 2). They then generated several more or less original responses as to what this scenario could represent, such as “candle in room” (“low original”) and “jelly fish in ocean” (“high original”). Following the task, participants self-rated the originality of their responses, as done in the prior research. Researchers analyzed the EEG data to identify the profile of brain waves (i.e., the neural signature) as participants generated high versus low original responses.

An additional feature of both experiments was that they were designed to assess whether the alpha acoustic beats produced an effect on brain waves that persisted during a one-minute baseline EEG recording immediately following their presentation.

CONCLUSIONS

CASL’s research on the neural signature of divergent thinking supports the following conclusions:

1. Engaging the left hemisphere more than the right provides a neural signature for good divergent thinking.

The neural signature of good divergent thinking, particularly of the ability to generate highly original responses, consists of increased alpha brain waves in the right hemisphere. This result replicates previous findings.

Further, the statistical test results enabled a more refined description of this neural signature than was provided in previous EEG studies. Highly original responses are associated with a larger amount of alpha brain waves in the right than left hemisphere, whereas there is no such difference for less original responses.

Greater engagement of the left than right hemisphere in active, effortful processing may underlie this effect. Active, effortful processing is indicated by a decrease in alpha brain waves. The left hemisphere may have been involved in more active, effortful processing during the verbal thinking task because it governs speech and language and because the task involved generation of verbal solutions under time pressure.

2. Alpha acoustic beats provide a useful method for determining and interpreting brain wave patterns associated with good divergent thinking.

The alpha acoustic beats provided a useful method for identifying and interpreting the neural signature of good divergent thinking. During divergent thinking, alpha brain waves were lower in the left than right hemisphere in the control condition, when alpha acoustic beats were not present, but they did not differ in the experimental conditions, when alpha acoustic beats were present.

The alpha acoustic beats—which were designed to increase alpha brain waves and create an alert but calm state—may have prevented the left hemisphere from increasing its active, effortful processing and reduced the fluency and originality of divergent thinking. Support for this interpretation comes from the finding that withholding of the alpha acoustic beats increased the fluency (i.e., number of ideas expressed) of divergent thinking and from the finding that the left hemisphere was more engaged than the right hemisphere when generating highly original ideas.
3 Alpha acoustic beats do not improve verbal divergent thinking, at least not when there is a relatively strict limit on problem solving time.

The experimental procedure used in the present investigation provided only a limited time for thinking of answers to each test item. The left hemisphere may have had to work with greater cognitive effort under such time constraints, especially since it is the hemisphere that is dominant for speech and language and since the divergent thinking task involved generation of verbal responses. The alpha acoustic beats may have prevented the left hemisphere from becoming more actively engaged by increasing alpha brain waves, which reflect an alert but calm and receptive (i.e., noneffortful) mode of thinking.

RELEVANCE

The findings are of both theoretical and practical relevance. The theoretical relevance is that the obtained neural signature of good divergent thinking (i.e., greater left than right hemisphere engagement) casts doubt on the frequently made claim that the brain’s right hemisphere is more crucial than the left hemisphere for the creative aspects of cognitive task performance. As suggested above, engagement of the left hemisphere may be more crucial than engagement of the right hemisphere for fluency and originality in divergent thinking, particularly when the task is a verbal one and when effortful responding within a short time limit is required. The practical relevance is that the findings suggest that it would be worthwhile to identify a method that can quickly increase the engagement of the left hemisphere during a divergent thinking task. Such a method, especially if it is easy to adopt, could then be inserted in a course aimed at training the divergent thinking of analysts. An activity that requires no training and that is as simple as squeezing a rubber ball in one’s right hand for just 45 seconds, for example, may increase the activation level of the left hemisphere.

CONSIDERATIONS

Examining the role of relaxed processing: Although the present study concluded that effortful engagement of the left hemisphere is important for promoting good divergent thinking, it is also possible that a more relaxed mode of processing of the left (and right) hemisphere might benefit divergent thinking. Such a relaxed mode of processing could be especially beneficial when it comes after a period of intense, effortful attempts at solving a problem, since it may allow the problem solver to overcome mental fixation and gain access to solution paths that are initially not considered. Elucidating the conditions and task phases under which effortful versus relaxed modes of thinking benefit the problem solving of analysts is of considerable interest.

Listening to pink noise to improve divergent thinking: In this study, we found some indication that one minute of listening to pink noise only while not engaged in a cognitive task increased alpha brain waves, compared to when acoustic beats were presented as well. It might be worthwhile to examine whether one minute of listening to pink noise only increases alpha brain waves compared to a condition in which no pink noise is presented. If it does increase alpha brain waves, then listening to pink noise may provide a simple method for inducing a relaxed receptive state of mind that might benefit divergent thinking during certain conditions or task phases (e.g., at the end of a period of effortful work on a problem when there is temporary cessation of flow of new ideas and solutions).

ENDNOTES

11 Pink noise is similar to white noise, but in pink noise, the lower frequencies are emphasized.

Corresponding Author and Reprints: Henrik J. Haarmann, PhD, Area Director, Cognitive Neuroscience, University of Maryland Center for Advanced Study of Language, (301) 228-8826, hhaarmann@casl.umd.edu.

Funding/Support: This material is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this report is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the authors of the product, process, or service that is the subject of this report. No one may use any information contained or based on this report in advertisements or promotional materials related to any company product, process, or service or in support of other commercial purposes. This report is not Releaseable to the Defense Technical Information Center per DoD Directive 3200.12.

The Contracting Officer’s Representative for this project is David Cox, Government Technical Director at CASL, (301) 226-8970, dcox@casl.umd.edu.