Training Auditors to Think Skeptically: Experiment 2

David Plumlee  
David Eccles School of Business  
University of Utah, Salt Lake City, UT 84112  
david.plumlee@business.utah.edu

Brett A. Rixom  
David Eccles School of Business  
University of Utah, Salt Lake City, UT 84112  
brett.rixom@business.utah.edu

Andrew J. Rosman  
University of Connecticut School of Business  
University of Connecticut, Storrs, CT 06268-1041  
Andy.Rosman@business.uconn.edu

Draft; do not quote without permission. Comments welcome.

April 2012

This research was funded by a grant from the Center for Audit Quality. However, the views expressed in this article and its content are those of the authors alone and not those of the Center for Audit Quality. We thank Margot Cell from the Center for coordinating the participation of auditors from the Big Four firms in the training modules and assessments. The comments made by Mark Nelson, Eldar Maksymov, Bart Ward and workshop participants from the University of Oklahoma, University of Utah, and University of Nevada, Las Vegas, on earlier versions of the manuscript are appreciated. We also thank Erik Boyle and Kreg Delaney for their many hours spent coding data.
Experiment 2

In the first experiment, we included three of four possible combinations of divergent and convergent thinking training (both, neither and divergent-only) and found that training in both improves diagnostic reasoning better than divergent training alone. Our design choice leaves open the question of whether convergent thinking requires training in divergent thinking to be effective. In formulating H2, we maintain that the typical or default mode of generating explanations involves a continuous winnowing of explanations known as “consistency checking” (Gettys and Fisher 1979, Fisher et al. 1983, Gettys et al.1986). H2 predicted that auditors trained to employ a sequence of divergent followed by convergent thinking will not resort to “consistency checking,” and as a result will generate a larger initial set of explanations during divergent thinking due to a reduced concern toward their initial plausibility. While those trained in both skills generated more explanations than either the divergent thinking only or the control group, the first experiment did not offer evidence about the use of consistency checking by individuals do not receive training in both. Thus, we design and conduct a second experiment to examine the efficacy of training in convergent thinking alone and gather data on the process individuals employ in generating explanations.

The second experiment incorporates two training conditions: full training (both divergent and convergent thinking) and convergent-only. As in experiment 1, both conditions received a total of four training modules whose content, except for the exclusion of some of the knowledge and comprehension questions, was virtually identical to the materials used in Experiment 1. In terms of design, all participants received the same materials for modules 1, 3, and 4: the introduction, convergent thinking training, and the final assessment, respectively. For participants in the full training condition the second module focused on how to use divergent
thinking, while the second module for the convergent-only condition contained one of the alternative tasks from the first experiment. They participated via the same online survey management software previously used and the other procedures were the same, except participants were given 24 hours to complete each of the first three modules and 48 hours for the last one.

The participants were forty-six Masters of Accounting students from a large public university enrolled in a fraud and forensic accounting class who were randomly assigned to one of the two conditions. They had taken an average of 2.33 audit-related courses and the two conditions differed slightly in the number of these courses taken (full, 2.55 vs. convergent, 2.13; \( t = 1.86, \text{prob.} < 0.070 \)). The participants had worked an average of 0.60 “busy seasons,” but did not differ in the number of “busy seasons” worked (full, 0.55 vs. convergent, 0.65; \( t = -0.29 \text{prob.} < 0.772 \)). The experimental conditions did not differ in their motivation (\( t = 1.05, \text{prob.} < 0.299 \)), general auditing knowledge (\( t = 0.71, \text{prob.} < 0.48 \)) or comprehension of the experimental materials (\( t = 0.51, \text{prob.} < 0.612 \)).

The analysis for the second experiment focuses on the questions of whether training in convergent thinking requires divergent thinking training to be effective, and on the nature of the process by which alternatives are generated. Based on the same detailed directions the coders in the first experiment used, two of the authors who were blind to the participant’s condition independently coded the written explanations from the final assessment case as well as the explanation chosen as the most likely cause. They substantially agreed on the appropriate coding (initial agreement was 90.9%; \( p < 0.000 \)), and all disagreements were reconciled.

---

1 One participant indicated he/she was in the tax-track of the Masters of Accounting program with limited auditing knowledge and was excluded from all analyses. His/her exclusion had no effect on the inferences drawn from the results.
The average number of unique explanations for the full-training condition was 4.36 and 3.30 for the convergent-only condition, which were significantly different in the expected direction \((t = 1.865, p < 0.04, \text{ one-tailed})\). The difference between the conditions was effectively the same as the full and divergent-only conditions in the first experiment, though the condition averages were lower in the second experiment. We attribute this latter finding to the participants’ less-developed knowledge structures in the second experiment. Their limited audit experience meant that they had less audit knowledge to tap into for potential explanations. Four of the 45 participants identified one of the two possible correct explanations, and all four of these were in the full training condition, which is significantly higher than the convergent-only condition \((\chi^2 = 4.59, p = 0.03)\) and is consistent with our expectation. Similar to the explanations generated, it is a lower proportion than in the first experiment. Again, we believe this was caused by the participants’ less-developed knowledge structures due to their limited audit experience. These results reinforce and extend those from the first experiment. We find that convergent thinking training alone is neither sufficient in terms for the number of explanations generated nor does in improve the ability to combine available information to identify specific causes.

To elicit the role of “consistency checking,” we included four questions about the process participants used when they generated explanations in the final assessment case. They were included in the materials following participant’s generation of explanations and selection of their most likely explanation. The questions along with the participants’ average responses on a seven-point Likert scale are shown in Table 1. The average responses to three of the four questions were significantly different, while the difference for a fourth question was moderately significant; all in the expected direction. In addition, we combined these four questions into a scale of “consistency checking use” (Cronbach’s Alpha = .738) and find the two conditions
differed significantly in the expected direction (full, 5.64 vs. convergent, 3.99, \( t = 4.79, p < 0.000 \), one-tailed). These results strongly support our contention that training in both divergent and convergent encourages anyone engaged in the diagnostic reasoning process to generate more explanations than those with convergent training only. Training in both leads individuals to consciously keep explanations they generate for later evaluation during the convergent thinking phase of the diagnostic reasoning process. Without training in both, they spontaneously engage in consistency checking where they evaluate explanations as they occur eliminating some based on superficial consideration.
Table 1

Participants’ responses in the second experiment to four questions that assess their use of consistency checking when generating alternative explanations in the assessment case.

<table>
<thead>
<tr>
<th></th>
<th>Full n = 22</th>
<th>Convergent n = 23</th>
<th>t stat</th>
<th>prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>My list includes all of the explanations that I could think of including explanations that I thought may not hold up under closer scrutiny.</td>
<td>5.45</td>
<td>4.70</td>
<td>1.302</td>
<td>0.100</td>
</tr>
<tr>
<td>I mentally checked the explanations as I thought of them to see if they made sense and I eliminated those that didn't make sense. (Numbers have been reverse coded.)</td>
<td>4.86</td>
<td>2.74</td>
<td>5.190</td>
<td>0.000</td>
</tr>
<tr>
<td>I consciously kept explanations on my list knowing that I could choose not to pursue them.</td>
<td>5.73</td>
<td>3.87</td>
<td>3.810</td>
<td>0.000</td>
</tr>
<tr>
<td>I think it's okay that some of the explanations on my list might not hold up under closer examination.</td>
<td>6.50</td>
<td>4.65</td>
<td>3.527</td>
<td>0.001</td>
</tr>
<tr>
<td>All four questions as a “consistency checking” scale.</td>
<td>5.64</td>
<td>3.99</td>
<td>4.786</td>
<td>0.000</td>
</tr>
</tbody>
</table>