

Thinking Inside the Box: The ShadowBox Method for Cognitive Skill Development

Gary KLEIN^a, Neil HINTZE^b, and David SAAB^b

^aMacroCognition LLC

^bBattalion Chief, FDNY

^bMacroCognition LLC

ABSTRACT

One way to help trainees develop expertise is to let them see the world through the eyes of experts. However, the tasks of gaining access to the expert's cognition and then of making experts available for training are daunting and impractical. Recently, however, Hintze (2008) developed a technique to allow trainees to shadow the thinking of experts. The trainees work through scenario-based materials, entering their information and decision priorities in a series of one-inch-square boxes. At the end of each scenario the trainees calibrate their interpretations with the conclusion of a panel of experts. Thus, the method does not require the a subject-matter expert to facilitate the training. Hintze used this ShadowBox method to increase expertise of firefighters, and it is currently being applied to a DARPA project for developing social interaction skills.

KEYWORDS

Training; scenario-based training; expertise.

INTRODUCTION

One way to help trainees develop expertise is to let them see the world through the eyes of experts. This “expert view” would let trainees discover what experts think is important in a situation, how they focus their attention, and also what they ignore. It would help trainees broaden their viewpoints and appreciate how subtle events might have important implications.

Bloom and Broder (1950) provided this type of perspective in a project designed to help under-performing college students do better on multiple choice tests. Bloom and Broder collected think-aloud protocols from students who did very well on these tests. Then they had under-performing students also generate think-aloud protocols as they struggled with the same items. Next, Bloom and Broder showed each under-performing student his/her own transcript along with a transcript from a successful student, and asked, “What was the difference?” For example, some under-performing students noticed that when they didn't know the answer they gave up, whereas the successful student shifted from a recall/recognition mode into a problem solving mode, trying to figure out what the answer might be, or at least trying to eliminate a few of the options. Bloom and Broder did not offer any advice. They let the under-performing students make their own discoveries about how they were falling short and what they needed to do on future tests. Their method generated significant improvements in performance. This study illustrates the impact of letting trainees see how experts (or people more skilled) view a task. The under-performing students defined deficiencies that were meaningful to them, which helped them translate training requirements into personal action.

Hintze (2008) provided a similar opportunity for firefighters. He developed scenario-based exercises coupled with cognitive task analysis materials to allow newly promoted officers to see the scenarios through the eyes of experienced officers. Hintze developed four challenging scenarios about unusual emergency situations. The scenarios were presented in a booklet and given to 14 experienced New York City Fire Department Officers who described how they would handle the decisions and wrote down the rationale for their decisions. These data were synthesized to illustrate the “expert mindset” to the experimental group.

These scenario materials were presented to an experimental group of 14 recently promoted New York State Fire Department Officers. The matched control group consisted of 15 New York City Fire Department Officers. The training materials exposed novices to situations they would not experience during routine operations. The training increased decision making performance by 18% ($p < .001$) for the experimental group over the control group. Hintze created a scoring key with 100 points as the maximum. The average score for the experimental group was 86.9, compared to 73.6 for the controls.

The Hintze study demonstrates a means of providing trainees with access to the thinking process of experts without having the experts present. It provides a means of enabling trainees to compare their responses to those



Authors retain copyright
of their work

H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.). *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013), Marseille, France, 21-24 May 2013*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-00-1

of the expert panel and determine for themselves what the differences are, and to speculate about the limits of their decision making.

Subsequent to this study, this methodology has been used to train additional groups of firefighters, and it has evolved into the ShadowBox method described below.

PRACTICE INNOVATION

The ShadowBox method is scenario-based. The trainees receive a booklet presenting a challenging scenario. They also have a second booklet for recording their answers at predetermined decision points. They record their answer to questions posed at each decision point in a separate small box, usually one-inch-square (but sometimes two-inches square). The trainees are essentially trying to shadow the expert panel by seeing the match between the responses they enter in the answer boxes to the answers provided by the expert panel.

Thus, for a given scenario, the initial page might describe the immediate situation including a map or photograph. At the bottom of the page is a prompt to enter any information they want to remember in the box for decision point #1, along with their rationale for what they included in the box. Once the trainees finish (they are given about 2.5 minutes for this task), they can never turn back in the booklet. All they will have to go on is what they wrote down:

- (1) The scenario passage is read aloud to the group, they follow along.
- (2) When the facilitator finishes reading, the trainees record what they want to remember in the ShadowBox.
- (3) The trainees then answer the decision point questions.
- (4) The facilitator then compares their Shadowbox entries and decision point answers to that of the experts, explaining the rationale for experts' answers.
- (5) The trainees compare their responses to those of the panel, and compare the rationale. They are asked to describe the differences in the contents of the two boxes.
- (6) The facilitator moves on to the next scenario update.

There are several forms of Shadow box. First, with the Attention box, trainees select what to put in the Shadow box to remember. Second, at some of the decision points the trainees have an Action Priority box; the booklet lists a small set of potential actions and the trainees prioritize these in order of importance and enter the top three into this second type of box, along with their rationale in the margin of the page. Then they learn the priorities listed by the experts, along with their rationale. Again, they have an opportunity to compare their responses and rationale to those of the experts.

A third type of box is the Information box. For some of the decision points the trainees are instructed to enter one query into the Information box—one type of information they would like to have at this point in the scenario or a question they would like to pose to the expert panel, along with their rationale for asking this question. Then they learn what the experts wrote, and go through the same comparison.

Other types of boxes are possible. There can be an Anticipation box (what is likely to happen in the next 15 minutes), an Assessment box (asking about different possible explanations of what is happening, often in form of yes/no questions), a Monitoring box (which cues should be watched most carefully) and so forth.

One of the most labor-intensive activities for implementing the ShadowBox method is to obtain responses from an expert panel. In his Master's thesis, Hintze (2008) used 14 FDNY officers as Subject Matter Experts (SME). He interviewed all of them face-to-face. Because of scheduling limitations, Hintze interviewed some of the experts 1:1, and others in small groups of 5-8 experts. These small groups were more efficient to run, and also provided valuable dialog that helped to improve the responses. The disagreements prodded the small groups to define their rationale more clearly.

Based on these interviews with experts, an answer key is prepared. The answer key consists of the consensus response for what should go into each box, along with a summary of the rationale responses. Hintze found a strong consensus for many of the boxes, but never achieved 100% convergence. In some cases the experts did not reach a strong consensus, and Hintze let the trainees know about any strong minority position that had emerged. He made it clear that there was no ground truth for any of the answers.

PILOT TEST

We applied the ShadowBox method to the task of training social skills in police officers. This pilot was part of a DARPA program, "Strategic Social Interaction Modules (SSIM)." The purpose of the SSIM program was to understand why some police and military personnel are more skilled at interacting with civilians to gain voluntary compliance, and to turn the findings into training to build interpersonal skills in police and military. The nickname for the program is the "Good Strangers" program because the intent is to develop skills of positive interactions with civilians rather than relying on intimidation to gain compliance.

Three scenarios were developed, using incidents that had been probed through the use of a Critical Decision method interview. The scenarios revolved around a domestic violence incident, an incident of managing a gang during the funeral of one of their members, and a case of persuading a suicidal man to drop his weapons. We prepared ShadowBox forms for each incident, using a mix of different types of queries. We collected calibration

data using a panel of seven Subject Matter Experts interviewed separately by telephone. We conducted an initial pilot test with seven participants to modify the materials.

We collected pilot data on 16 experienced police officers working in the Spokane Washington area. Their experience level ranged from 7 to 29 years. We ran four groups of four officers. It took approximately four hours to run each group. We had two facilitators (neither of them with any police experience) guiding the groups. As the week continued, we added a wrinkle of having the officers fill out the priorities from the perspective of a “Bad Stranger,” or from the perspective of a rookie. We broadened the Attention box to include inferences and questions, not just information to remember.

We collected evaluation data from 15 of the sixteen police officers. We asked them to provide ratings using a 5-point scale where 5 was high and 1 was low. When asked if the scenarios were realistic, the police evaluated them as a 4.3. When asked if the scenarios were interesting and engaging, the rating was 4.4. The question “I learned a lot from this exercise” was rated 3.8. For the question of whether the training should be delivered by an experienced police officer, seven said yes and seven said no (there were two non-responders). When asked if they would do anything differently having gone through the scenarios, the responses ranged from 50% to 87.5% who said that they would do something different, for the three scenarios.

To illustrate, one of the ShadowBox scenarios we used for the Good Strangers project involved a case of domestic violence. The actual incident was a bit unusual in that several police squad cars arrived on the scene, and when the senior officer entered the home he had three other police officers behind him. In the scenario (as in the actual incident), the senior officer gained entry to the home after knocking and getting no response. Upon entry the officer saw an agitated and frightened eight-year-old boy opened the door. The boy was the one who called 911. Sounds of a fight were clear – emanating from a bedroom on the second floor. The scenario stopped at this point to inject a decision point: what are the priorities?

The choices were: (1) Ignore the boy and focus on the screams; (2) Inform the boy why the police are there and calm him down before moving on to the threat; (3) Direct the other officers to handle the threat upstairs (4) Remove the boy to stairs outside the apartment. The scenario continued to the arrest of the abusive boyfriend, the entry of the bloodied mother who becomes outraged to learn that her son had summoned the police, and so forth.

The later decision points addressed alternative strategies for handling the mother’s anger at her son and abuse of the police, and the criteria for deciding when it was safe to leave – without endangering the son who might be physically punished by his out-of-control mother. For the purpose of this paper, we will only address the first decision point. Before reading further, you might consider how you would prioritize these options.

The panel of SMEs reviewed the four options and by a narrow margin selected option (3) as the top priority. The rationale including the following: the boy was a potential witness; the boy was a victim-by-proxy; this was an opportunity to demonstrate to the boy that police officers can be trusted to provide security; the boy is at risk in the future if the boyfriend returns; the boy’s needs should be respected; the large contingent of police officers made it unnecessary for all four to go charging up the stairs. The panel gave the lowest priority score to option (1) Ignore the boy and focus on the screams.

Of the 16 police officers tested on this Decision Point, zero selected (2) as the top priority. Five of them made it their lowest priority. The ShadowBox exercise was thus very effective for confronting these police officers with the mismatch between their own tendencies and the perspective of a Good Stranger.

DISCUSSION

The method permits trainees to “shadow” experts by seeing how their own responses compared to the responses from the expert panel, and to speculate, using the rationale materials, about how the experts were thinking. For his original research, Hintze estimated that trainees needed 2 to 2.5 hours to complete the four scenarios, about 30-40 minutes per scenario. The trainees had about 2.5 minutes to fill in the box for each decision point. Hintze was the facilitator for the training sessions for his Master’s thesis. He read the scenario aloud as the trainees followed along with their booklets, pausing at each decision point to let the trainees fill in the boxes for the questions posed. We needed 4 hours for our three scenarios.

The ShadowBox method is similar to techniques such as Tactical Decision Games (TDGs, Schmitt, 1996) and Decision-Making Exercises (DMXs, Klein, 2005) that present scenarios and invite trainees to respond. The ShadowBox method seems to have some advantages over these other approaches in that it provides an answer key generated by an expert panel (although, like the TDGs and DMXs it does not claim that there is any “right” answer), it provides the rationale for the experts’ answers, and then it provides a convenient format for trainees to explicitly provide their responses. The contrast between trainee answers/rationale with those of the expert panel helps trainees appreciate the limits of their mental models and points the way for them to develop their cognitive skills further.

When we applied the ShadowBox method to the DARPA Good Strangers project, one skill we addressed was to take someone else’s perspective, particularly when the other person comes from an unfamiliar culture. Here, the ShadowBox method was similar to “cultural assimilators,” a method that has been validated in the field of intercultural communication. The method of cultural assimilators relies on a collection of real-life scenarios describing puzzling cross-cultural interactions and explanations for avoiding the emerging misunderstandings.

This is one of the most researched and accepted methods of cross-cultural training (Bhawuk & Brislin, 2000; Landis & Bhagat, 1996, Albert, 1983).

The ShadowBox method differs from techniques such as the U.S Army's "Think Like a Commander" program (Lussier & Shadrack, 2002; Ross & Lussier, 1999) which tries to develop cognitive skills via tactical scenarios. The Think Like a Commander program has defined a core set of cognitive skills for the military officer, such as keeping a focus on the mission and the commander's intent, modeling a thinking enemy, considering effects of terrain, using all assets available, considering timing, seeing the big picture, visualizing the battlefield, and remaining flexible. The training emphasizes these principles. In contrast, the ShadowBox method does not pre-define any principles. The training is all contextualized, and the trainee learns by peering into the minds of the experts—seeing their priorities and responses and the rationale they offer. In addition, the ShadowBox method can be used with non-kinetic missions such as counterinsurgency, which are different than the kinetic missions and core skills the Think Like a Commander program was designed for.

One of the important potential extensions of the ShadowBox method is to adapt the materials so that they do not require expert facilitation. In theory, the ShadowBox method could be in an on-line form without any facilitator. Participants could drag and drop items from the scenario into the answer box, speeding up the process and reducing a source of ambiguity.

Challenges for the ShadowBox method are to determine whether the scenarios used are representative of the domain to be handled, and to develop efficient means for knowledge capture with domain experts. Other extensions would be to use video representations of scenarios rather than written/text representations, and to use systematic scoring methods to enable computer-based scoring for immediate feedback. The ShadowBox format is very flexible; in addition to a training application it can be adapted for knowledge capture for a cognitive task analysis effort.

ACKNOWLEDGMENTS

This work was supported by The Defense Advanced Research Projects Agency (government contract/grant number 06-1825383). The views, opinions, and/or findings contained in this (article, poster, paper etc.) are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense. We gratefully acknowledge the assistance provided by Lois James, Washington State University.

REFERENCES

- Albert, R. D. (1983). The intercultural sensitizer or culture assimilator. A cognitive approach. In D. Landis & R. W. Brislin (Eds.), (Vol. 2, pp. 186-217). New York: Pergamon.
- Bhawuk, D. P. A., & Brislin, R. W. (2000). Cross-cultural training: A review. *Applied Psychology: An International Review*, 49(1), 162-191.
- Bloom, B. S., & Broder, L. J. (1950). *Problem-solving processes of college students: An exploratory investigation*. Chicago: University of Chicago Press.
- Hintze, N. R. (2008). *First responder problem solving and decision making in today's asymmetrical environment*. Unpublished Master's thesis, Naval Postgraduate School, Monterey, CA.
- Klein, G. (2005). *The power of intuition*. New York: A Currency Book/Doubleday.
- Landis, D., & Bhagat, R. (Eds.) (1996). *Handbook of intercultural training*. Newbury Park, CA: Sage Publications.
- Lussier, J. W., & Shadrack, S. B. (2002). *Think Like A Commander: Captain's Edition - Prototype 1.0*: ARI Research Product 2003-01 available from U.S. Army Research Institute for the Behavioral and Social Sciences, 2423 Morande Street, Fort Knox, KY 40121.
- Ross, K. G., & Lussier, J. W. (1999). A training solution for adaptive battlefield performance *Proceedings of the 1999 Interservice/Industry Training, Simulation, and Education Conference*. Orlando, FL.
- Schmitt, J. F. (1996). Designing good TDGs. *Marine Corps Gazette*, 80(5), 96-97.