Evaluating Two Modes of Observational Learning in Cognitive-Spatial Task Training

Nirit Gavish[1] and Michal Shelef[2]

[1] Ort Braude College, Karmiel, Israel
[2] Technion – Israel Institute of Technology, Haifa, Israel
Nirit@braude.ac.il, shelefm@tx.technion.ac.il

Abstract. The focus of the current study was to evaluate the effect of two modes of observational learning, dyad trainer-trainee performance and preliminary observational learning part, on training, as well as the interaction between them. We conducted an experimental study with a 3-D computerized puzzle. Each trainer offered four trainees instruction in solving this puzzle in a 2X2 between-participants design: with or without preliminary observational learning, and with dyad trainer-trainee performance or with verbal guidance only during training (16 trainees in each group). Results demonstrated that the preliminary observational learning resulted in longer training time but better performance in terms of success rates, and that dyad trainer-trainee performance led to shorter training time and did not influence performance. No significant interaction between the two modes was found. The cost-effectiveness matrix that was found in this study can assist in designing guidelines for choosing the appropriate observational learning methods in training.

Keywords: Observational learning, Training, Dyad performance, Cognitive tasks.

1 Introduction

In active learning – the standard means of learning a psychomotor task – the performer learns by doing, acquiring competence by trial and error or practice. In observational learning, a person learns by watching the actions of another person and imitating them (e.g., in a woodworking or cookery class) or mentally recording their effects (e.g., watching the driver from the passenger seat; watching a football game on TV). During learning by observation, the learner is not physically engaged in performing the task, and so can engage his full mental energy on monitoring the other person's actions and responses. Given the great power of human imagination and inferential processes, this method of learning has the potential to be highly effective. Indeed, the value of observational learning has been demonstrated in many studies (e.g., [1-3]), and observational learning has been proved to produce good cognitive representations of skilled actions [4-5].

When an expert in the task serves as the trainer, taking advantage of observational learning may be even more worthwhile. First, observing an expert enables the learner

D. Harris (Ed.): EPCE/HCII 2013, Part I, LNAI 8019, pp. 222–231, 2013.

to build up an accurate representation of the task, which may be different from his or her intuitive representation; this representation allows the learner to imitate (either immediately, or later on) the necessary correct actions [3]. As Herrnstein, Loewenstein, Prelec and Vaughan [6] and Herrnstein and Prelec [7] pointed out, in a multiple-alternative space, when the value of alternatives is not known, subjects will stop exploring as soon as they hit an acceptable response, and thus converge to a local optimum. Many complex tasks can be performed in many ways, so demonstrating the best way to complete the task can prevent this convergence. Second, performing a complex task can put a high cognitive load on the novice, preventing him from absorbing the necessary information. Observing an expert can release the cognitive resources needed for focusing on the essential elements of the task [2].

Observational learning can take various forms. One of them is dyad performance. In dyad performance, the trainer and trainee – or, alternatively, two trainees – perform the task together. A variant, triad training, involves either three learners or two learners and a trainer. Shebilske, Regian, Arthur and Jordan [8] tested dyad and performance protocols using a complex computer game called Space Fortress [9]. In the dyad condition, trainees practiced in pairs, each controlling half of the task (performing half of its roles). Despite having only half as much hands-on experience during practice sessions, the dyad trainees performed as well as individual trainees on ten test sessions. Replications of this study have produced similar results of improving training efficiency by reducing training time without sacrificing performance (See [10-15]).

Observational learning can also be part of a two-stage learning process, where the learner first watches an expert performs the task, and only then engages in active practice. As described above, through observation the learner builds an accurate mental representation of the task, which he or she can then imitate during the active-learning phase [4]. As Wouters et al. [2] point out, such learning makes it less likely that novices will adopt inappropriate automated patterns of behavior that are difficult to change. This type of learning paradigm is also in keeping with the strategy of beginning with easier forms of training and progressing to more difficult stages [16] - a strategy widely employed in modern training.

Choosing between different methods to employ observational learning during training requires knowing the effects of each method on the training and the unsupervised performance following it. Employing more than one method requires also being aware of the interaction between the selected methods. The current study explored the use of the abovementioned two modes of observational learning: dyad performance and preliminary observational learning part. The aim was to identify the separate effects of each of these methods on training and performance and the interaction between the two methods. Based on past research on dyad performance protocol [9-15], it was predicted that dyad trainer-trainee performance will reduce training time while performance time will not be affected. It was also predicted, based on research on the benefit of observational learning at the beginning of the training [2], that preliminary observational learning, in which the trainee is usually given enhanced information on the task, will result in longer training time, but better performance. An additional prediction was that combining these two modes of observational learning will achieve the best result: better performance with shorter training time.

The task chosen for this study was a three-dimensional computerized puzzle task, used previously by Yuviler-Gavish et al. [17]. Training and performance in this task was found to be affected by employing different training methods. For example, Yuviler-Gavish et al. [17] demonstrated that training participants in this task with visual guidance in addition to verbal guidance was attractive to both trainers and trainees and demanded less cognitive effort from trainees, but at the same time impaired skill acquisition relative to verbal guidance only. It was hence assumed that the use of different observational learning strategies to train this task will produce significant effects on training and performance.

In the current study, training was performed on the computerized task. The trainers' goal was to teach trainees to solve a puzzle *on their own*, and trainees' performance was evaluated in non-supervised tests on both the computerized task and the real-world version of it. Combinations of two modes of observational learning were examined in a 2X2 design: the first mode - preliminary observational learning (with versus without,) and the second mode - dyad trainer-trainee performance (versus verbal guidance only).

2 Method

2.1 Design

The study used a 2X2 between-participants experimental design with randomized order, where the training variables were with preliminary observational learning (hereinafter, "Preliminary") versus without preliminary observational learning (hereinafter, "No Preliminary"), and with dyad trainer-trainee performance (hereinafter, "Dyad") versus verbal guidance only ("Verbal"). The four conditions were created by combining the four training variables described above. In condition Preliminary-Dyad, trainers demonstrated and verbally explained the task while the trainee watched and listened, until both trainer and trainee felt ready to move on to the active training phase. During the active training, trainers could continue to use both verbal guidance and pointing with the mouse or manipulating objects as needed. In condition Preliminary-Verbal, training in both phases was limited to verbal explanation only; trainers were not given access to a mouse and so could not demonstrate or point. In the two No Preliminary conditions, training began with the active training phase. In condition No Preliminary-Dyad, trainers were allowed to give verbal guidance and point with the mouse or manipulate objects, while in condition No Preliminary-Verbal, they could give verbal guidance only.

2.2 Participants

Ninety-five participants were recruited for the study. Nineteen (12 males and 7 females) received extensive training in the 3-D computerized puzzle task and served as trainers. The 76 trainees, 39 males and 37 females, were evenly distributed among the four experimental conditions (9 males and 10 females in the No Preliminary–Dyad condition, 10 males and 9 females in each of the other conditions). Each trainer trained four trainees, one in each condition, in random order. All participants were undergraduate or graduate students at the Technion - Israel Institute of Technology. Participants' average age was 24.2 (ranging from 20 to 31). None of the trainees had experience with the specific task used in the experiment, and only 17% had any experience with computerized puzzles.

Due to a technical recording problem, training data from three trainees (two of them were trained by the same trainer) and real-world test data from an additional trainee were not available for this analysis. In order to maintain the trainer's repeated-measures analysis design, all other trainees' data belong to these four trainers were not included in the analysis (12 altogether). Hence, the analysis includes only the data of sixteen trainees, 9 males and 7 females, and sixty-four trainees, 30 males and 34 females (8 males and 8 females in the No Preliminary-Verbal condition, 7 males and 9 females in each of the other conditions).

Participants were paid a fixed amount of NIS 40 (about USD 10) per hour. Additionally, trainees could receive a bonus according to their performance in the computerized test phase (either NIS 50 or 100), and trainers received a bonus according to their trainees' performance in this test (ranging from NIS 5 to 60 for each trainee).

2.3 Experimental Task and Setup

The main task required participants to solve a three-dimensional puzzle, called Shuzzle (*http://www.leweyg.com/lc/shuzzle.html;* see Figure 1). In this puzzle, pieces must be moved in a virtual three-dimensional space so as to complete a shape indicated by a wire frame. Horizontal movements are performed by dragging the pieces with the left mouse button. Vertical movements involve dragging the pieces with the left mouse button while pressing a key. Vertical rotation is achieved by dragging with the right mouse button, and horizontal rotation by dragging with the right mouse button while pressing a key.

Each trainer sat in a booth next to his trainee, to enable vocal communication between them and avoid any kind of non-verbal communication (e.g., gestures) that was not tested in this study. A white rectangular plastic divider between the booths prevented trainers and trainees from seeing each other. The trainee used a regular PC with two 19 inch screens, one in the trainee's booth and one in the trainer's booth, so that both trainer and trainee had the same visual presentation. A second computer mouse attached to the computer extended into the trainer's booth, allowing the trainers in the Dyad condition to use the mouse for pointing and manipulating objects.

The test phase also included a real-world version of the puzzle. In the real puzzle, colored pieces must be arranged so as to create a given shape, indicated by a hollow metal model (see Figure 2). The pieces were made from lightweight polypropylene (Delrin); each piece weighed about 40-60 grams.

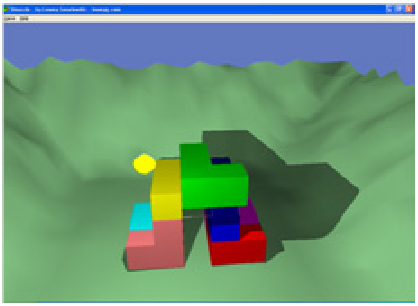

Fig. 1. Top: The computerized 3-D puzzle task. Bottom: The two practice tasks.

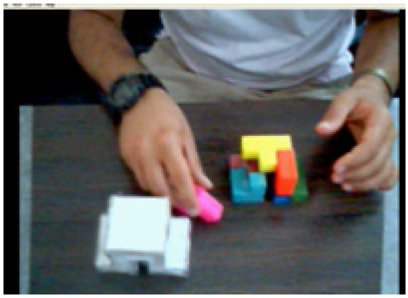

Fig. 2. The real-world puzzle task

2.4 Procedure

The experiment had five stages: an initial practice stage; training on the computerized puzzle; two tests, one on the computerized puzzle and one on the real-world version of it; and a subjective evaluation questionnaire.

Participants in all four conditions were first given a brief explanation of the puzzle objective and general information on how to perform the various movements (as noted above). Then, participants completed an initial practice stage consisting of two easy puzzles with two or three shapes (see Figure 1, bottom). During the practice stage, participants could ask questions and the experimenter was permitted to remind participants of the appropriate keys to perform each routine. This stage was included so that the main challenge in the training phase would be the assembly of the three-dimensional object rather than the details of how to handle the puzzle pieces.

Following the initial practice, trainees were informed that their goal in the experiment would be to complete the same puzzle as quickly as possible *on their own* following a training phase. They were told that their payoff would depend on their completion time in the computerized test: The best performer in their respective condition (of 19 individuals) would receive a bonus of NIS 100, and the four runners-up would each get a bonus of NIS 50. The experimental training phase followed. Trainees were then given the two tests – first the same computerized puzzle task they had trained in (Figure 1, top) and then the real-world puzzle version of it (Figure 2).

Participants recruited as trainers gathered at the lab the day before the main experiment. Following the initial practice stage, they performed the main puzzle task on their own three times. Nineteen participants who succeeded in the main puzzle task were invited back the next day. At that point, they were informed that their goal in the experiment would be to train four novice participants, and that they would be rewarded based on the success of the four participants in a non-supervised test. The trainers were instructed about the training protocol for each condition. Each trainer trained one participant from each group, in random order.

Questionnaires were completed after the tests by both trainers and trainees. Each questionnaire included four items to be answered on a scale from 1 to 5; trainers answered the questions separately for each of their trainees. For trainers, the questions were: 1) How easy was it for you to train this trainee? 2) How well did the trainee understand your training, in your opinion? 3) Do you think you succeeded in transmitting the message for the trainee in a clear fashion? 4) How difficult was the task for the trainee, in your opinion? For trainees, the questions were: 1) To what extent was the training understandable? 2) Do you think the trainer succeeded in transmitting the message in a clear fashion? 3) How much did the training help your task performance? 4) Rate the level of task difficulty. After completing the questionnaires, trainers and trainees were thanked, debriefed, and given their bonuses. Trainers' bonuses were based on their trainees' completion times in the computerized test in minutes (COMP) according to the following formula: BONUS = 60- 2.5×COMP, rounded up to the nearest multiple of 5.

3 Results

3.1 Training

Training times, as well as test times, were analyzed using repeated-measures ANOVA, with Preliminary observational learning (Preliminary or No Preliminary) and Dyad Performance/Verbal Guidance conditions as the within-participant variables. The unit of analysis for the repeated-measures tests was the trainer (data from his four trainees were analyzed as repeated-measures).

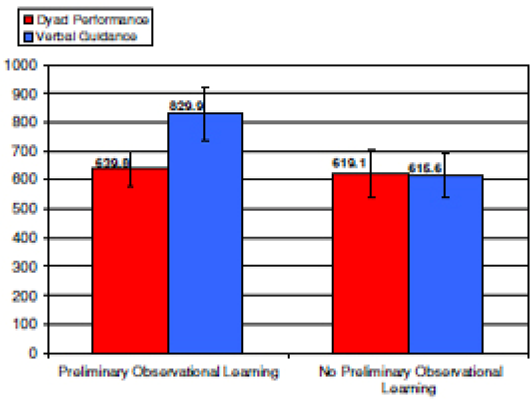

Fig. 3. Mean training time (in seconds) for the four experimental groups. The error bars denote the standard error.

Training time (including the time of the preliminary observational learning) was significantly longer for the Preliminary observational learning condition compared with the No Preliminary condition (Preliminary mean: 734.8 seconds, SD = 322.4; No Preliminary mean: 617.4 seconds, SD = 309.3; $F(1,15) = 5.32$, p = .036, *Partial Eta Squared* = 0.262). In contrast, training time in the Dyad Performance condition was significantly shorter than in the Verbal Guidance condition (Dyad mean: 629.4

seconds, *SD* = 282.3; Verbal mean: 722.8 seconds, *SD* = 350.1; $F(1,15) = 4.9$, $p = .043$, *Partial Eta Squared* = 0.246). The interaction between Preliminary and Dyad/Verbal was not significant ($F(1,15) = 2.09$, $p = .169$, *Partial Eta Squared* = 0.122). See Figure 3. The differences between training time levels in the experimental conditions for the 12 participants that were not included in the analysis were the same as for the analyzed data (showing longer training time for the Preliminary condition compared with the No Preliminary condition: a difference of 183.7 seconds, and shorter training time for the Dyad condition compared with the Verbal condition: a difference of 278.2 seconds).

3.2 Test Performance

Following the supervised training, participants were tested on the same computerized puzzle without the trainer's supervision, then on the real-world version. Failure in the computerized puzzle test was defined as not managing to successfully solve the puzzle within 30 minutes. In the real-world puzzle this time was reduced to 10 minutes, because this test did not include difficult computerized manipulations in order to rotate the shapes or to move them, and thus it could be performed (and was performed) much faster. Success rates were analyzed using Multinomial Logistic Regression with Preliminary observational learning (Preliminary or No Preliminary) and Dyad performance/Verbal Guidance conditions as factors.

For the computerized puzzle test, success rates were significantly higher for the Preliminary observational learning condition compared with the No Preliminary condition (Preliminary mean: 100%; No Preliminary mean: 90.6%, $\chi 2(1, N = 64) = 4.33$, $p = .038$). For the Dyad and Verbal conditions, however, they did not differ significantly (96.9% and 93.8%, respectively; $\chi 2(1, N = 64) = 0.37$, $p = .541$). See Figure 4. The mean performance time for the Preliminary observational learning condition compared with the No Preliminary condition was not significantly different, probably due to the high variability among participants in the No Preliminary condition (Preliminary mean: 193.3 seconds, *SD* = 149.8; No Preliminary mean: 336.4 seconds, *SD* = 506.6; $F(1,15) = 2.95$, $p = 1.107$, *Partial Eta Squared* = 0.164). Performance time in the Dyad Performance condition was similar to that of the Verbal Guidance condition (Dyad mean: 237.0 seconds, *SD* = 342.8; Verbal mean: 292.8 seconds, *SD* = 412.9; $F(1,15) = 0.30$, $p = .593$, *Partial Eta Squared* = 0.020). The interaction between Preliminary and Dyad/Verbal was not significant ($F(1,15) < 0.01$, $p = .974$, *Partial Eta Squared* < 0.001). See Figure 5.

No significant effects were found for the real-world test. Success rates were not significantly different for the Preliminary observational learning condition compared with the No Preliminary condition (Preliminary mean: 87.5%; No Preliminary mean: 90.6%, $\chi 2(1, N = 64) = 0.17$, $p = .685$), and also not significantly different for the Dyad and Verbal conditions (84.4% and 93.8%, respectively; $\chi 2(1, N = 64) = 1.49$, $p = .222$). The mean performance time for the Preliminary observational learning condition was not significantly different compared with the No Preliminary condition (Preliminary mean: 115.7 seconds, *SD* = 191.1; No Preliminary mean: 97.7 seconds, *SD* = 190.9; $F(1,15) = 0.15$, $p = 0.707$, *Partial Eta Squared* = 0.010). Performance time in the Dyad Performance condition was also not significantly different compared with the Verbal Guidance condition (Dyad mean: 148.2 seconds, *SD* = 223.5; Verbal mean: 97.7 seconds, *SD* = 147.25; *F*(1,15) = 0.96, *p* = .344, *Partial Eta Squared* = 0.060). The interaction between Preliminary and Dyad/Verbal was not significant, either (*F*(1,15) = 0.08, *p* = .781, *Partial Eta Squared* = 0.005).

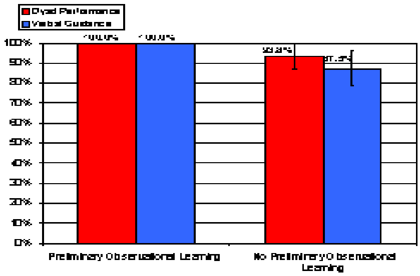

Fig. 4. Success rates in the computerized test for the four experimental groups. The error bars denote the standard error.

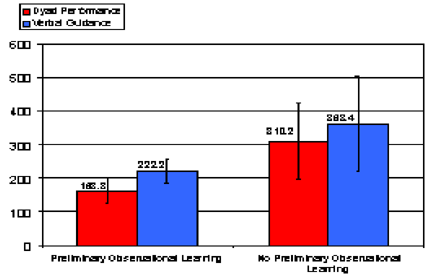

Fig. 5. Mean test time in the computerized test (in seconds) for the four experimental groups. The error bars denote the standard error.

3.3 Subjective Evaluation of the Training

In the questionnaires completed after the tests, three significant effects of training conditions were found. Trainers reported that the they succeeded more in transmitting the message, Question 3 for trainers, in the Dyad Performance condition (mean score 1.78 out of 5, *SD* = 0.49) than in the Verbal Guidance condition (mean score 2.31, *SD* = 1.15, *F*(1,15) = 10.87, *p* = .005, *Partial Eta Squared* = 0.420), and also rated the difficulty of the task for the trainee, Question 4 for trainers, as lower in the Dyad condition (mean score 2.56, *SD* = 0.88) than in the Verbal condition (mean score 3.03, *SD* = 1.15, *F*(1,15) = 7.83, *p* = .015, *Partial Eta Squared* = 0.343). In addition, trainees felt that the training helped their task performance, Question 3 for trainees, more in the Preliminary condition (mean score 1.09, *SD* = 0.30) than in the No Preliminary condition (mean score 1.38, *SD* = 0.55, *F*(1,15) = 7.64, *p* = .014, *Partial Eta Squared* = 0.338).

4 Discussion

Observational learning, in which a person learns how to perform a task by watching another person perform the relevant steps, has been shown to be effective in many experimental tasks as well as in everyday life (e.g., [1-3]). It has also been shown that observational learning can produce good cognitive representations of skilled actions [4-5]. Using an expert in the task as a trainer can increase the value of observational learning even more.

Two different modes of observational learning, in an expert-trains-a-novice paradigm, were examined in this study in a 2X2 between-participants design for a computerized 3-D puzzle task. In the first mode, dyad performance, the trainer and trainee perform the task together during the training process (see [8]). In the second, preliminary observational learning, the training starts with a demonstration, and the trainee performs the task only after acquiring a basic knowledge of the necessary steps (see [2], [16]). It was predicted that: Dyad trainer-trainee performance will reduce training time and will have no effect on performance; Preliminary observational learning will result in longer training time but better performance; And that the combination of them will achieve better performance with shorter training time.

Results demonstrated that the preliminary observational learning resulted in longer training time but better performance (in terms of success rates), and that dyad trainer-trainee performance led to shorter training time and did not influence performance, in line with our assumption. However, no significant interaction between the two modes was found. Participants' subjective evaluations were in favor of these two modes. The lack of significant findings for the real-world test in the present research can be a consequence of the preceding computerized test, which eliminated possible differences between conditions. Trainers reported that they succeeded more in transmitting the message in the dyad performance condition, and also rated the task as easier on this condition. Trainees reported that the training helped their task performance more in the Preliminary condition than in the No Preliminary condition.

The findings of this study can assist in designing guidelines for choosing the appropriate observational learning methods in training. A cost-effectiveness matrix was found, in which preliminary observational learning costs additional training time, but improves performance. In contrast, dyad trainer-trainee performance saves training time without sacrificing performance. Designers of training protocols and training systems can choose between these two methods based on training goals. The lack of interaction between the two methods may be due to the current study's limitation or a robust finding, and it is recommended that future research will focus on ways to best combine the two methods. In addition, the transfer of learning to similar tasks should be further evaluated.

The task chosen for this study is a cognitive-spatial task, in which the motor elements are of lesser importance. In addition, the training included a preliminary stage meant to overcome the major motor difficulties that trainees may encounter. Findings from the current study may be limited to this specific kind of tasks, and should be further evaluated using tasks with different characteristics. Another shortcoming of the selected task is that once the solution for the puzzle was discovered, it is easy to perform the necessary steps, and as a result a relatively large variance across trainees was produced, especially for the condition without preliminary observational learning. Taking advantage of a task with a smaller expected variance may ease to reach statistical significance even for the factors which were not found to have significant effect in the current study.

Acknowledgment. This research was supported in part by the European Commission Integrated Project IP-SKILLS-35005.

References

1. Bandura, A., Jeffery, R.W.: Role of Symbolic Coding and Rehearsal Processes in Observational Learning. Journal of Personality and Social Psychology 26, 122–130 (1973)
2. Wouters, P., Tabbers, H.K., Pass, F.: Interactivity in Video-based Models. Educational Psychology Review 19, 327–342 (2007)
3. Wulf, G., Shea, C.H.: Principles Derived from the Study of Simple Skills Do Not Generalize to Complex Skill Learning. Psychonomic Bulletin & Review 9, 185–211 (2002)
4. Bandura, A.: Social Foundation oF Thought and Action: A Social Cognitive Theory. Prentice-Hall, Englewood Cliffs (1986)
5. Carroll, W.R., Bandura, A.: Representational guidance of action production in observational guidance of action production in observational learning: A causal analysis. Journal of Motor Behavior 22, 85–97 (1990)
6. Herrnstein, R.J., Loewenstein, G.F., Prelec, D., Vaughan Jr., W.: Utility maximization and melioration: Internalities in individual choice. Journal of Behavioral Decision Making 6, 149–185 (1993)
7. Herrnstein, R.J., Prelec, D.: Melioration: A theory of distributed choice. Journal of Economic Perspectives 5, 137–156 (1991)
8. Shebilske, W.L., Regian, J.W., Arthur, W., Jordan, J.A.: A dyadic protocol for training complex skills. Human Factors 34, 369–374 (1992)
9. Mane, A., Donchin, E.: The Space Fortress game. Acta Psychologica 71, 17–22 (1989)
10. Arthur, W., Strong, M.H., Jordan, J.A., Williamson, J.E., Shebilske, W.L., Regain, J.W.: Visual attention: Individual differences in training and predicting complex task performance. Acta Psychologica 88, 3–23 (1995)
11. Arthur, W., Day, E.A., Bennett, W., McNelly, T.L., Jordan, J.A.: Dyadic versus individual training protocols: Loss and reacquisition of a complex skill. Journal of Applied Psychology 82, 783–791 (1997)
12. Arthur, W., Young, B., Jordan, J.A., Shebilske, W.L.: Effectiveness of individual and dyadic training protocols: The influence of trainee interaction anxiety. Human Factors 38, 79–86 (1996)
13. Day, E.A., Arthur, W., Shebilske, W.L.: Ability determinants of complex skill acquisition: Effects of training protocol. Acta Psychologica 97, 145–165 (1997)
14. Sanchez-Ku, M.L., Arthur, W.: A dyadic protocol for training complex skills. Human Factors 42, 512–520 (2000)
15. Shebilske, W.L., Jordan, J.A., Goettl, B.P., Paulus, L.E.: Observation versus hands-on practice of complex skills in dyadic, triadic, and tetradic training-teams. Human Factors 40, 525–540 (1998)
16. Yechiam, E., Erev, I., Parush, A.: Easy first steps and their implication to the use of a mouse-based and a script-based strategy. Journal of Experimental Psychology: Applied 10, 89–96 (2004)
17. Yuviler-Gavish, N., Yechiam, E., Kallai, A.: Learning in multimodal training: Visual guidance can be both appealing and disadvantageous in spatial tasks. International Journal of Human-Computer Studies 69, 113–122 (2011)