

# The Effect of Voice Instruction on the Construction of Mental Model

Restyandito<sup>1</sup>, Alan H.S. Chan<sup>1</sup>, and Umi Proboyekti<sup>2</sup>

<sup>1</sup> Dept. of Systems Eng. & Engineering Mgt., City University of Hong Kong,  
Hong Kong SAR  
rrestyand2@mslive.cityu.edu.hk, alan.chan@cityu.edu.hk

<sup>2</sup> Dept. of Information Technology, Duta Wacana Christian Univ.,  
Yogyakarta, Indonesia  
othie@ukdw.ac.id

**Abstract.** The goal of this study is to observe the effect of instruction deliverance method in the construction of mental model. A good mental model can help the user's learnability process. There were two methods tested in this study: a step-by-step instruction (SS) and a complete set of whole-steps instructions (WS) to finish a given task. The SS group performed better on the learning process, however they had the least score on both the information retention and transfer process. Their minds were not engaged in the process, as they seemed to simply follow the instructions without being critical. When error occurred, they tended to be less persistent in trying to finish the task. This might be caused by the incomplete mental model as a result of receiving the instruction step by step.

**Keywords:** voice instruction, mental model, learnability.

## 1 Introduction

The use of speech technology as an interface to access the internet has been increasing in the past decades [1-4]. It furnishes new chances to enhance the accessibility by providing compensation for limitations of specific user groups. [5]. Speech is also often used in a multimodal interface to augment communication between human and machine. In their user study, Hofmann et al. found that users are willing to use and trust a speech dialog system [2]. Speech as an input method also has advantages such as simpler, faster and more convenient to use than other methods [6]. However, Boufardea et al. [7] pointed out speech interface can never be perfect. There are some factors that can influence the effectiveness of spoken language interface such as acoustics, speaking style, out-of-vocabulary words and understanding gaps as well as technological limitation [5].

The goal of this study is to study the effect of given voice instruction in the construction of mental model. Mental model is defined as the knowledge that the user has about how a system works, its component parts, the processes, their interrelations, and how one component influences another [8]. It is useful for learning, information

retrieving and problem solving [9-11]. This study will focus on the use of mental model in the learnability process. Jay Forrester commented “*A mental model changes with time and even during the flow of a single conversation*” (as cited in [12]). Therefore the way an instruction is delivered will have different impacts on how users build and construct their mental models, which will also influence their user experience and effectiveness of the interaction in the learning process.

## **2 Previous Study**

### **2.1 Auditory Interface**

The continued development of technologies and their application open up demands and opportunities toward multimodal interface. Many studies have been conducted on the effectiveness of multimodal interfaces [5],[6],[13]. A study by Rigas et al. [14] suggested that interfaces could be designed in a way that visual metaphors communicate the information that ‘needs’ to be conveyed to the user and the auditory metaphors (earcons) communicate the other part of information (the interaction part) which is used to perform tasks. However, the use of auditory representations must consider certain factors. Speech is language dependent and often too slow, pure tones are easily confused with each other, while musical instrumentation though easier to listen to it needs learning and abstraction because of the intuitive mapping [15]. It was found that the use of combinations of auditory icons, earcons, speech, and special sound effects helped user to make fewer mistakes in accomplishing their tasks, and in some cases reduced the time taken to complete them [13]. An experiment by Sodnik et al. [16] showed that auditory interfaces were effective but were not faster than the visual interface to use in a mobile environment. Auditory interfaces can be beneficial to reduce the cognitive workload when visual perception is needed to attend to other task, such as driving, or operating machine. Burke et al. [17] conducted a meta-analysis on the effects of multimodal feedback on user performance. They found visual and auditory cue provide advantages in reducing reaction times and improving performance scores, but it does not reduce error rates effectively. Furthermore it was found that visual-auditory feedback was most effective in single task scenarios, because the use of both auditory and visual channels will increased user’s workload hence less advantageous in situations where high workload condition are already present. This study will focus only on the auditory interface in the form of voice instruction.

### **2.2 Mental Model in Learning**

Visual imagery has been known as one of the techniques to improve human memory [18]. Greek and Roman orators use this technic to keep track of the many parts of their long speeches. They visualize an imaginary place and as they walk through this place in their mind, they attach objects associated to the parts of their long speeches. When they need to recite their speech, they just need to visualize and walk through that place in their mind [19].

Several studies on the role of mental model in learning to operate a device showed that people who used mental model learned the procedure faster, retained it more accurately and executed them faster [8],[10],[20]. Other studies on the impact of visualization on the learning process have also been conducted. The work of Scwamborn et al. [21] indicated positive main effects of learner-generated pictures on drawing and mental effort on comprehension measurement. Students may learn better from text and pictures than from text alone, because pictures increase appropriate active processing during learning while reducing non relevant cognitive processing. Leutner et al. [23] conducted an experiment to 10th grade students, where they were asked to mentally imagine text content while reading an expository science text. They found that mental imagery increases text comprehension, even though visualization strategies could cause high demands of cognitive load on the learner [22]. They also pointed out “*Decreased cognitive load due to constructing mental images has, in terms of a main effect, no direct impact on reading comprehension and thus, on learning*”.

Phillips et al. [11] stated “*Instructions are a common example of communicating models of technological systems and can act as a boundary object between designers’ conceptual models and models developed by users*”. When delivering the instruction, it is significant to keep the wording of signs and instructions as simple and short as possible [5]. The instruction must be able to communicate a clear decision pathway [24].

### 3 Method

#### 3.1 Participants

Seventy five first year university Indonesian students (60 male, 15 female) participated in this study. Their mean age was 18.74 years (SD=1.64). All participants received monetary expense allowance for their time. They were familiar with internet and have some experience in using search engine. Table 1 provides a summary of the participants’ background.

**Table 1.** Respondent background experience

		n	%
Gender	Male	60	80.00
	Female	15	20.00
Time spent on the internet daily	< 5 hours	28	37.33
	5-10 hours	32	42.67
	> 10 hours	15	20.00
Using search engine to find information	Always	64	85.33
	Sometimes	11	14.67
	Never	0	0
Using attribute in search engine	Often	9	12.00
	Sometimes	55	73.33
	Never	11	14.67
Using logic operator in search engine	Often	5	6.67
	Sometimes	36	48.00
	Never	34	45.33

### 3.2 Materials

The experiment was conducted in a university laboratory with intranet access to the student project report repository website. The software used to capture the screen movement was Snagit ver 11.2.0.101. The browser used to access the repository website was Firefox ver 26.0. The computer specification for this experiment was Intel Core 2 CPU 4300 @ 1.80 GHz, with 2014 MB of RAM.

### 3.3 Design

There are two ways of giving instructions, namely explicit demand and implicit demand [2]. This study will investigate two approaches of delivering instructions: a short one step at a time instruction, and a long instruction explaining the whole series of steps. Participants were randomly assigned to one of the three treatment conditions with 25 participants in each group. The first group is the control group (CG) which received no instruction, the second group received a short step by step instruction (SS), and the last group received long instructions describing the whole set of steps (WS). To minimize technological limitation which may cause ineffectiveness in user's perception as pointed out by Neerincx et al. [5], the instruction was delivered live by a person using decent sound system instead of the speech synthesis and pre-recorded messages. The instruction given was clear and using direct vocabulary by take into consideration what the participants knows [24]. All participants were familiar in using the internet and should have some experience of using the search engine. For this between-participant experiment the participants were asked to find some information from the final year project report repository web site of undergraduate students at Duta Wacana Christian University.

Before the experiment, participants received a brief introduction about the experiment which was to study the effect of mental model in the learnability process. The participants from SS and WS were asked to construct mental model in their mind as they listened to the instruction. Participants were assured that it is ok if they could not complete the task and the result would be anonymous. They also received short training on how to use Snagit to record their activities during the experiment.

Figure 1 shows a snapshot of the final year project report repository website. Participants can search information by typing the keyword directly on the search bar ("Budi Susanto"), or refine their search in several ways, such as using attribute (dozen1: "Budi Susanto"), click on the categorical result on the left side (grouped by Department, research topic, and supervisor), or scroll manually using scroll bar.

There were four tasks in the experiment. The first task was to find out how many projects were supervised by *Gloria Virginia* on *Genetic Algorithm*. This task can be completed first by typing the attribute (*supervisor1*) and the keyword (*Gloria Virginia*) on the search bar, and further refine the result with the title of the project (*Genetic Algorithm*). The first task is the learning process in which participants received instructions on how to retrieve that information using attribute. The control group (CG) had to ascertain information by trial and error. The second task was to find supervisors who supervised projects on *Multi Criteria* by *Ferry Himawan*. This time, no instructions

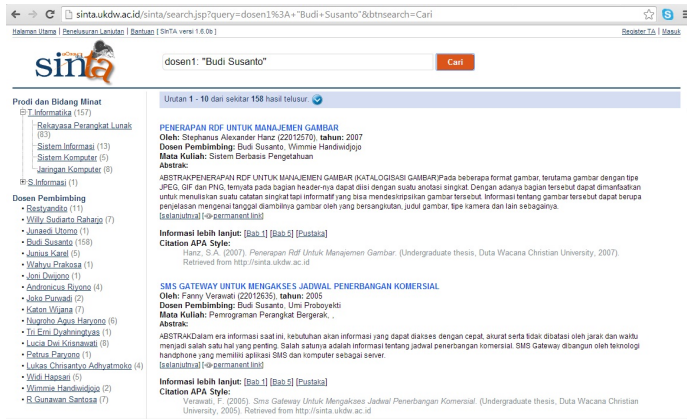


Fig. 1. Sinta, the final year project report repository website

were given to all of the groups. Participants were expected to rely on their previous experience to finish the task. This process is known as the retention process. Participant could retrieve the information using the same step they did on the first task, but with different attribute (*title*) and keyword (*Multi Criteria*) and further refined the result by the student’s name (*Ferry Himawan*). For the third task, participants were asked to find information on the abstract of projects supervised by *Budi Susanto* on the topic of *resource description framework*. The purpose of this task was to learn the usage of logic operator. Participants could retrieve the information by using ‘AND’ operator in addition to search attribute. Similar to the first task, they received instructions on how to do it except that the control group received no instruction at all. The last task was aimed to examine the transfer process of using previous knowledge and experience to carry out task in new condition different from the previous tasks. They did not receive any instruction. Participants were asked to find a project based on incomplete information. They were only given the first name of the supervisor (*Umi*), they did not have information whether *Umi* is the first supervisor or the second supervisor, and they only knew the topic of the project instead of the title. Hence they needed to make use of the operator and attribute in slightly different ways. They had to use three attributes (*supervisor1*, *supervisor2*, *topic*) and both operators (OR and AND) in the search bar. The purpose of the second and fourth task is to discover how mental model help participants in completing similar task and new task. For each task, the number of errors (NE), success rate (SR), and completion time (CT) were recorded. The performance of each group will be compared.

## 4 Results and Discussion

Table 2 presents the summary result of the experiment. A one way between subjects ANOVA was conducted to compare the effect of instruction delivery on success rate, average number of errors and average completion time. There was a significant effect on instruction delivery methods on success rate only on Task 1 [F(2,72)=10.031,

$p = 0.0001$ ]. There was a significant effect on instruction delivery methods on completion time, on Task 1 [ $F(2,72)=4.227$ ,  $p = 0.019$ ] and Task 3 [ $F(2,72)=3.597$ ,  $p = 0.034$ ]. There was also a significant effect on number of error, on Task 1 [ $F(2,72)=4.777$ ,  $p = 0.012$ ] and Task 3 [ $F(2,72)=7.064$ ,  $p = 0.002$ ]. However, there was no significant effect on instruction delivery methods on success rate, completion time and number of error on Task 2 and Task 4.

**Table 2.** Summary of participant's performance for each task

<b>T1</b>	CG	SS	WS	<b>T2</b>	CG	SS	WS
Success Rate (SR)	4	20	10	Success Rate (SR)	19	14	20
Avg. Num. of Error (NE)	1.88	2.02	2.72	Avg. Num. of Error (NE)	2.84	2.18	2.72
Avg. Compl. Time (CT)	288s	212s	253s	Avg. Compl. Time (CT)	156s	184s	156s
<b>T3</b>	CG	SS	WS	<b>T4</b>	CG	SS	WS
Success Rate (SR)	19	21	21	Success Rate (SR)	3	1	6
Avg. Num. of Error (NE)	0.80	2.06	2.52	Avg. Num. of Error (NE)	4.80	4.56	4.24
Avg. Compl. Time (CT)	146s	192s	173s	Avg. Compl. Time (CT)	276s	293s	286s

Task 2 was meant to see the retention process based on the experience of completing Task 1. The results here show the group which received Step-by-Step instruction (SS) had the lowest performance, while the best performance was achieved by the group which received Whole-Steps instructions (WS). The same performance can be seen on the transfer process in Task 4, where SS had the lowest SR and WS achieved highest SR. The SS participants might not be able to build the mental model of the overall system when they received the step by step instructions, while the WS participants had to construct the mental model as they received the series of instructions before executing them. Hence WS participants had a more complete mental model which they can explore to finish a task. A complete series of instructions enable participants to visualize a decision pathway in their mind [24]. However, this observation must be tested further, as the ANOVA analysis for this experiment did not show statistical difference.

As expected, there is an increase in the success rate of Task 3 compared to Task 1. This might be the result of the learnability effect [9] from the previous two tasks. Both Control Group (CG) and WS show a meaningful improvement on SR, 475% and 210% respectively. While for SS, improvement in SR was noted from one participant only (105%). Based on the observation of recorded screen activities, there were some factors contributing to the failure of following instructions such as spelling error, preconception based on past experience or knowledge, interpretation and familiarity. An example of preconception is when the participants were asked to type *title*: “*genetic algorithm*”, one of the participants was typing *title*= “*genetic algorithm*”, which might be influenced by his/her experience using query language. Other examples of mistakes caused by interpretation are when participants were asked to type in the search bar *title*: “*multicriteria*” AND *supervisor*: “*budi susanto*”, there were several

interpretations of the instruction, one of participant typed *title: "multicriteria" & supervisor: "budi susanto"*, other participant typed *title: "multicriteria" N supervisor: "budi susanto"*. There was a participant who did not type the query in the search bar of the site, instead he typed it in the search bar of the browser. An example of mistakes caused by familiarity was when the participants were asked to search a final project supervised by Mrs.Umi of the Management Department, there were some participants who automatically typed Mrs.Umi Proboyekti who is a lecturer at the Information Technology Department where the participants came from and familiar with. Neerinx et al. [5] recommended the use of instruction as short as possible, however this might lead to ambiguity. In this example punctuation mark matters, as well as case sensitive. The instruction given, did not explicitly inform participants how to type the searching query, resulting some of them made mistakes in the first try. Yet, giving very detail instruction might not be effective either, as the instruction will be long, and human's memory ability is limited [18]. Too much detail instruction may cause participants to forget what they need to do [25]. These examples show the challenges of making a clear instruction as there are many factors that can influence users' comprehension [7].

Step-by-Step instruction did not seem to stimulate users to think actively as they listen to the instruction; hence they just typed whatever they thought they had heard. Furthermore, when they encountered an error, they were less likely to try to solve the problem. As seen in Table 2, they had less number of errors compared to the WS participants, even though they did not perform better on completing the task. This could indicate their lack of determination to finish the task. They might conclude that the instruction given to them was wrong.

The performance of all groups was worst for Task 4. This might be caused by the fact that participants needed to transfer their knowledge to complete a new task. Had they more exposure using the system, they might perform better. Supportive information is essential in the process of acquiring cognitive skills, but so is practice [26], whether those skills are recurrent (performed the same way on each incidence) or non-recurrent (performed differently according to conditions managed by complex rules or contextual features) [27].

It can also be seen in Table 2 that the high number of errors does not necessarily mean low success rate. For example, even though number of errors for the CG and WS groups were higher than SS, it turned out they also had higher success rate. Users were versatile and got adapted to the system operation easily. If they ran into errors when using an application, they would naturally try to solve them, by adapting, improvising or negotiating [6]. Most likely participants would have to rely on their mental model to predict the step needed to complete the task. Marchionini [28] found that the efficiency of mental model building depends on the level of detail transferred. Table 3 shows the summary of the regression analysis based on the instruction delivery method and past knowledge or experience in using the search engine. The R Square value is 6%-23% indicating there might be more variables influencing participants' performances, such as age, gender, intelligence level, etc.

**Table 3.** Summary of regression analysis

		SUCCES RATE (SR)							
		T1		T2		T3		T4	
		<i>Coeff.</i>	<i>P-val</i>	<i>Coeff.</i>	<i>P-val</i>	<i>Coeff.</i>	<i>P-val</i>	<i>Coeff.</i>	<i>P-val</i>
Step by Step	(SS)	0.276	0.070	-0.172	0.235	-0.024	0.870	-0.107	0.232
Whole Steps	(WS)	0.050	0.746	0.153	0.301	0.123	0.417	0.065	0.482
	Search Engine	0.221	0.268	0.080	0.675	0.011	0.953	-0.123	0.299
	Attribute	-0.055	0.631	0.046	0.679	-0.148	0.192	0.025	0.715
	Logic	-0.084	0.498	0.215	0.073	0.177	0.149	-0.089	0.229
		R Square = 0.139		R Square = 0.166		R Square = 0.070		R Square = 0.089	
		R = 0.372		R = 0.407		R = 0.264		R = 0.298	
		AVERAGE COMPLETION TIME (CT)							
		T1		T2		T3		T4	
		<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>
Step by Step	(SS)	-16.094	0.041	26.284	0.245	45.376	0.016	12.18	0.164
Whole Steps	(WS)	12.945	0.107	-6.347	0.783	22.792	0.229	4.769	0.593
	Search Engine	-9.186	0.371	-7.271	0.807	18.243	0.454	21.07	0.071
	Attribute	7.802	0.191	-13.383	0.438	-8.309	0.555	-1.201	0.857
	Logic	1.981	0.756	-7.199	0.698	-13.75	0.366	1.238	0.863
		R Square = 0.238		R Square = 0.067		R Square = 0.150		R Square = 0.097	
		R = 0.487		R = 0.258		R = 0.380		R = 0.311	
		AVERAGE NUMBER OF ERRORS (NE)							
		T1		T2		T3		T4	
		<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>	<i>Coeff.</i>	<i>P-value</i>
Step by Step	(SS)	1.273	0.032	-0.309	0.702	2.510	0.001	0.023	0.976
Whole Steps	(WS)	1.582	0.010	-0.242	0.771	1.708	0.018	-0.955	0.227
	Search Engine	0.730	0.346	-0.557	0.603	-0.431	0.636	-0.977	0.337
	Attribute	0.234	0.600	-0.767	0.218	0.073	0.890	0.656	0.266
	Logic	0.400	0.406	0.086	0.897	0.259	0.648	-0.171	0.786
		R Square = 0.178		R Square = 0.042		R Square = 0.207		R Square = 0.082	
		R = 0.422		R = 0.206*		R = 0.455		R = 0.286	

The minimum correlation coefficient R for 70 subjects and more with 95% confidence level is at least 0.232. There is only one instance where R is < 0.232, on the



average number of errors (NE) of Task 2. Consequently, it can be concluded that these variables contribute to the performance of the participants. The SR of Task 1 was affected mostly by the instruction delivery method (SS) as it has the highest coefficient of 0.276, while Task 2 and Task 3 were affected mostly by previous experience using Logic (0.215 and 0.177 respectively) while Task 4 was affected mostly by the instruction method (WS) with coefficient of 0.065. Prior experience can help participants to understand the system better. They could construct a mental model based on their experience in using search engine, and use the same reasoning to complete the task. However, having prior experience may lead them not to pay attention to the instruction. There were three participants who used the advance search facility, even though they were instructed to type manually the attribute and keywords on the search bar.

The participants of this study were mostly first year students, while the tasks given were related to finding final year project information usually needed by final year students. Therefore, they might not be really determined to complete the task as if they were final year students. However they were purposely invited to participate in this study, because as first year students, they might not have any prior experience in using the repository system.

## 5 Conclusion

This study has tried to look at the effect of given voice instruction in the construction of mental model, especially during the retention and transfer process. One way ANOVA analysis has showed that instruction delivery method plays an important role in participants' performance during the learning process. There was evidence that whole steps instruction enables participants to perform better during the retention and transfer process, although the result was not statistically significant.

Further study can be conducted with more participants and task related to participants' interest or need. The tasks given to the participants can be concretized using the Bollywood Method, where a task will be given a dramatic and exaggerated back-story to excite the participants into believing the urgency of the problem [29]. The study can also consider more variables for inclusion in the experiment.

Based on the findings, some recommendation can be made in the interface design using voice instruction.

- Instruction should be clear and precise to avoid ambiguity.
- Instruction given should fit user knowledge and experience
- Step-by-Step instruction ensure higher success rate in the learning process
- Instruction should not be too long, so user will not forget the instruction

**Acknowledgment.** The first author would like to thank the Directorate General of Higher Education of Indonesia for supporting his study at City University of Hong Kong.

## References

1. Lin, D., Bigin, L., Bao-zong, Y.: Using chinese spoken-language access to the WWW. In: 5th International Conference on Signal Processing, Proceedings of the WCCC-ICSP 2000, vol. 2, pp. 1321–1324 (2000)
2. Hofmann, H., Ehrlich, U., Berton, A., Minker, W.: Speech interaction with the internet - A user study. In: Proc - Int. Conf. Intelligent Environ., pp. 323–326 (2012)
3. Sazbon, M., Haddad, Y.: Advanced vocal web browser. In: IEEE Conv. Electr. Electron Eng. Israel, pp. 732–735 (2010)
4. Werner, S., Wolff, M., Eichner, M., Hoffmann, R.: Integrating speech enabled services in a web-based e-learning environment. In: Internat. Conf. Inf. Tech. Coding Comput., pp. 303–307 (2004)
5. Neerinx, M.A., Cremers, A.H.M., Kessens, J.M., van Leeuwen, D.A., Truong, K.P.: Attuning speech-enabled interfaces to user and context for inclusive design: Technology, methodology and practice. *Univers Access Inf. Soc.* 8(2), 109–122 (2009)
6. Knudsen, L.E., Holone, H.: A multimodal approach to accessible web content on smart-phones. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) ICCHP 2012, Part II. LNCS, vol. 7383, pp. 1–8. Springer, Heidelberg (2012)
7. Boufardea, E., Garofalakis, J., Plessas, A.: A dynamic voice portal for delivery of cultural content. In: Proc. - Int. Conf. Internet Web Appl. Serv., pp. 186–191 (2008)
8. Fein, R.M., Olson, G.M., Olson, J.S.: A mental model can help with learning to operate a complex device, pp. 157–158. *ACM* (1993)
9. Rouse, W.B., Morris, N.M.: On looking into the black box: Prospects and limits in the search for mental models. *Psychol. Bull.* 100(3), 349 (1986)
10. Kieras, D.E., Bovair, S.: The role of a mental model in learning to operate a device. *Cogn Sci.* 8(3), 255–273 (1984)
11. Phillips, R., Lockton, D., Baurley, S., Silve, S.: Making instructions for others: Exploring mental models through a simple exercise. *Interactions* 20(5), 74–79 (2013)
12. Smith, R.: Game impact theory: The five forces that are driving the adoption of game technologies within multiple established industries. *Games and Society Yearbook* (2007)
13. Rigas, D.I., Alsuraihi, M.: A toolkit for multimodal interface design: An empirical investigation. In: Jacko, J.A. (ed.) *HCI 2007*. LNCS, vol. 4552, pp. 196–205. Springer, Heidelberg (2007)
14. Rigas, D., Yu, H., Klearhou, K., Mistry, S.: Designing information systems with audio-visual synergy: Empirical results of browsing E-mail data, pp. 960-7620 (2001)
15. Wersényi, G.: Auditory representations of a graphical user interface for a better human-computer interaction. In: Ystad, S., Aramaki, M., Kronland-Martinet, R., Jensen, K. (eds.) *CMMR/ICAD 2009*. LNCS, vol. 5954, pp. 80–102. Springer, Heidelberg (2010)
16. Sodnik, J., Dicke, C., Tomažic, S., Billinghamurst, M.: A user study of auditory versus visual interfaces for use while driving. *Int. J. Hum. Comput. Stud.* 66(5), 318–332 (2008)
17. Burke, J.L., Prewett, M.S., Gray, A.A., Yang, L., Stilson, F.R.B., Coovert, M.D., Elliot, L.R., Redden, E.: Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis, pp. 108–117 (2008)
18. Baddeley, A.D.: *Your memory: A user's guide*. Sidgwick & Jackson, London (1982)
19. Carpenter, S., Huffman, K.: *Visualizing psychology*. Wiley. com (2009)
20. Ujita, H., Yokota, T., Tanikawa, N., Mutoh, K.: Computer aided instruction systems for plant operators. *Int. J. Hum. Comput. Stud.* 45(4), 397–412 (1996)

21. Schwaborn, A., Thillmann, H., Opfermann, M., Leutner, D.: Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Comput. Hum. Behav.* 27(1), 89–93 (2011)
22. Ainsworth, S.: The functions of multiple representations. *Comput. Educ.* 33(2-3), 131–152 (1999)
23. Leutner, D., Leopold, C., Sumfleth, E.: Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Comput. Hum. Behav.* 25(2), 284–289 (2009)
24. Harvey, G.: Designing procedural instructions: 5 key components. *Inf. Des. J.* 16(1), 19–24 (2008)
25. Deo, S., Nichols, D.M., Cunningham, S.J., Witten, I.H., Trujillo, M.F.: Digital library access for illiterate users (2004)
26. Darabi, A.A., Sikorski, E.G., Nelson, D.W., Palanki, S.: Efficient, motivational, and effective strategies for complex learning: Computer-based simulation experiments in troubleshooting. *Technology Instruction Cognition and Learning* 3(3/4), 233 (2006)
27. Darabi, A.A., Nelson, D.W., Seel, N.M.: The role of supportive information in the development and progression of mental models. In: *Learning and Instruction in the Digital Age*, pp. 101–115. Springer (2010)
28. Marchionini, G.: Making the transition from print to electronic encyclopaedias: Adaptation of mental models. *Int. J. Man Mach. Stud.* 30(6), 591–618 (1989)
29. Sherwani, J., Palijo, S., Mirza, S., Ahmed, T., Ali, N., Rosenfeld, R.: Speech vs. touch-tone: Telephony interfaces for information access by low literate users. In: *Int. Conf. Inf. Commun. Technol. Development*, pp. 447–457 (2009)