Title

Speaker/Gender Effect: Impact of the Speaker's Gender on Learning with Narrated Animations

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN

1069-7977

Authors

Gerjets, Peter
Linek, Stephanie B.
Scheiter, Katharina

Publication Date

2006

Peer reviewed

Speaker/Gender Effect: Impact of the Speaker's Gender on Learning with Narrated Animations

Stephanie B. Linek (s.linek@iwm-kmrc.de)
Hypermedia Research Unit, Knowledge Media Research Center,
Konrad-Adenauer-Str. 40, 72072 Tuebingen, Germany

Peter Gerjets (p.gerjets@iwm-kmrc.de)
Hypermedia Research Unit, Knowledge Media Research Center,
Konrad-Adenauer-Str. 40, 72072 Tuebingen, Germany

Katharina Scheiter (k.scheiter@iwm-kmrc.de)
Department of Applied Cognitive Psychology and Media Psychology, University of Tuebingen,
Konrad-Adenauer-Str. 40, 72072 Tuebingen, Germany

Abstract

Current cognitive multimedia design theories provide several guidelines for how to integrate verbal and pictorial information. Based on a dual-channel information processing model, many authors have suggested using spoken rather than written text along with pictorial representations (e.g., animations). However, the directives for the design of auditory text (narration) are still fragmentary, especially with regard to the kind of voice to be used. In the current paper a very fundamental question regarding the implemented voice was addressed, namely, whether to use a male or a female voice for the presentation of verbal information in multimedia learning environments. In the experiment reported learners studied animations that aimed at teaching students how to solve probability problems. The expository text was presented as narration. The learner's gender and the speaker's gender were varied between subjects. The results show that learners achieved better learning outcomes when the narration was presented by a female speaker rather than a male speaker. This finding occurred irrespectively of the learner's gender (speaker/gender effect). The results suggest augmenting purely cognitive approaches to multimedia design by social-motivational assumptions.

Cognitive Science and Multimedia Design

In the last ten years, multimedia learning environments have become ubiquitous means for conveying knowledge in the Social as well as in the Natural Sciences. Unfortunately in many cases the design of these environments is still driven by intuition and technological capabilities of the designers, rather than by empirically validated guidelines. However, recently theories of multimedia design based on Cognitive Science have been developed like Richard Mayer's (2001) cognitive theory of multimedia learning (CTML), from which guidelines regarding the design of verbal and pictorial representations in multimedia messages can be derived. The CTML states that incoming information consists of either verbal or pictorial representations, which are processed in a visual or a verbal channel depending on the modality of the representation (i.e., *dual-channel assumption*). According to the *limited capacity assumption*, these two separate channels are both limited with respect to the amount of information that can be processed at a time. The *active processing assumption* emphasizes that learners need to engage in sense-making activities and deeper cognitive processing of the instructional materials to achieve meaningful learning outcomes. Active processing includes the selection of relevant information, the organization of the selected information into coherent meaningful representations as well as the integration of the verbal and pictorial representations with each other and with prior knowledge (Mayer, 2001).

These assumptions have allowed deriving several design guidelines for multimedia learning, thereby inspiring numerous empirical investigations. But until now the focus of this research has mainly been on the question of how to combine pictorial and verbal information (e.g., temporal and spatial contiguity principle). However, comparably less work has been done to find out how each of the single representational formats should be best designed. While there are hardly any design recommendations for pictorial representations (e.g. Tversky, Morrison, & Betrancourt, 2002), some first insights exist with regard to how to present verbal information as part of multimedia messages. In the following section multimedia principles for the design of verbal information will be discussed.

Designing Verbal Explanations

According to the *modality principle*, dynamic visualizations that need verbal explanations are best presented as animations with auditory text (narrated animations) instead of written text (Moreno & Mayer, 1999). Theoretically, this principle has been ascribed to the dual-channel and limited capacity assumption. If written text is used to present explanations, the visual channel's limited capacity may be easily overloaded due to the fact that the verbal and pictorial information need to be processed within the same channel in parallel. On the other hand, presenting text as narration allows the distribution of processing requirements across both channels, the visual and the auditory channel, thus unburdening the cognitive system (The problematic issue of mapping the distinction between the two channels onto Baddeley's distinction between different working memory subsystems will not be discussed here, but is adressed in Rummer, Schweppe, Scheiter & Gerjets, submitted).

But when following the modality principle and using narration instead of written text, the question arises, which kind of voice should be used. This question goes beyond taking into account pure cognitive assumptions and pertains to additionally considering socio-motivational aspects in learning. However, there are also design principles for multimedia messages in CTML that rely on a socially enhanced view, for instance, the voice principle (Mayer, Sobko, & Mautone, 2003). The *voice principle* was investigated in a seminal study by Mayer, Sobko, and Mautone (2003), who compared the effectiveness of a standard accent versus foreign accent voice and a human versus a machine-synthesized voice that accompanied dynamic visualizations. They found that people learned better with a standard accent voice and with a human voice (voice principle).

Mayer et al. (2003) explained these findings by their social agency theory that is linked to CTML and suggests that people apply social rules to media, which in turn influences learning. The theory postulates several successive steps. First, it is assumed that stimuli like voices or pictures of the speaker in multimedia messages can act as *social cues*. According to this assumption, a human voice provides a stronger social cue than a machine-synthesized voice. Second, due to social cues, learners interpret the multimedia message as a kind of *social communication* in contrast to pure information delivery. Third, the interpretation of the multimedia learning scenario as a social communication situation leads to the activation of *social conversation schemas*. That is, social rules of human-to-human communication like the cooperation principle suggested by Grice (1975) are applied to the human-computer interaction. The social agency theory borrows this assumption from the media equation theory (Reeves & Nass, 1996) according to which persons tend to behave towards media as towards humans ("media equal real life", Reeves & Nass, 1996, p. 5). According to media equation theory that mainly describes affective and social-motivational variables "individuals' interactions with computers …are fundamentally social and natural" (Reeves & Nass, 1996, p. 5). The social agency theory by Mayer et al. (2003) ties these assumptions of media equation theory to cognitive aspects of learning. Fourth, following Grice's cooperation principle, learners assume that the speaker is trying to say something meaningful and therefore in turn try to make sense out of the spoken words. Social cues thus result in learners being more *motivated* and investing more *effort* to understand the spoken words. Accordingly, they engage in deeper cognitive processing of the instructional materials. Fifth, the result of such deeper cognitive processing is a *more meaningful mental representation*, which expresses itself mainly in better *transfer test scores*.

While this explanation of the voice principle offered by the social agency theory seems plausible at first sight, there are a couple of open questions with regard to the empirical evidence in favor of this explanation and with regard to the methodological approach taken towards the investigation of social-motivational aspects in multimedia learning. These issues are discussed in the next section by revisiting the voice principle and deriving consequences for future research.

The Voice Principle Revisited: Methodological Drawbacks and Consequences for Future Research

No Evidence for Mediators Most studies that rely on the social agency theory as a theoretical background have yet failed to provide support for the complex causal chain that is assumed by the theory. In particular, there is hardly any evidence for the assumption that motivation and effort mediate the relation between the presence of social cue and learning outcomes.

Methodical Variation of the Social Cue As Mayer et al. (2003) compared only two voices to each other, there may have been differences on more than just one dimension (i.e., accent). When comparing only *one* standard accent voice with *one* foreign accent voice it is rather impossible to keep all the other voice features like pitch, intonation, and other speaker-characteristics constant at the same time.

Alternative Explanations Based on the current empirical data it is yet unclear whether the social agency theory is needed for explaining the superiority of standard accent human voices over foreign accent human voices and machine-synthesized voices. Mayer et al. (2003) themselves admit that this pattern of results could also be explained by the cognitive load theory (CLT, Sweller, van Merriënboer, & Paas, 1998). According to the CLT explanation, processing a human standard accent voice imposes less cognitive demands onto learners, leaving more cognitive resources for deeper processing of the instructional materials. To rule out this explanation, it would be necessary to compare different voices with regard to cognitive load, but this measure has unfortunately not yet been obtained by Mayer et al. (2003).

Consequences for Future Research From these three methodological drawbacks, different consequences for future research can be drawn: First, a measurement of additional variables should be included, for instance, cognitive load during learning or potential moderators like the motivation to listen to different speakers or the mental effort invested. Second, multiple voices should be presented within each experimental condition so that differences can be traced back unambiguously to the experimental variation and are not caused by other speaker characteristics. Third, it would be more insightful to implement a comparison of voices that lead to different predictions depending on whether one takes a social-motivational or a purely cognitive perspective. As mentioned above, the voices compared by Mayer et al. (2003) lead to the same predictions from both perspectives. Thus, this comparison is not apt to decide whether the voice effects found can be best explained by social-motivational or purely cognitive mechanisms. Therefore, we suggest comparing voices that do not differ with regard to their cognitive processing demands, but that nevertheless activate different social schemas. These constraints are satisfied by varying the gender of the speaker. When contrasting male and female voices one would not expect any differences from a purely cognitive perspective, whereas a socially enhanced view allows for several predictions with regard to social-motivational variables, which in turn may affect learning. Therefore, speakers of different gender can serve as an appropriate experimental variation to clarify the relevance of social-motivational factors and give further insights into the relationship between cognitive and social aspects of multimedia learning.

According to the social agency theory, the gender of the speaker can have several implications for learning. For instance, it can be assumed that speakers of different gender activate different social schemas due to gender stereotyping, which is a ubiquitous phenomenon in human-human interaction (Franzoi, 1996). Research on media equation theory with regard to human-computer interaction has also demonstrated gender stereotyping towards media. For instance, computers with female voices were rated as being more competent regarding female topics like love and relationships, whereas male-voice computers were estimated as being more competent regarding male topics like mathematics or computers (Nass, Moon, & Green, 1997). Additionally there is also evidence for principles of interpersonal attraction towards media. For example, research on media equation reveals that computer users ascribe a personality to a computer based on verbal cues it delivers and prefer computers that resemble their own personality (Nass & Lee, 2001). However, there is not only evidence in favor of similarity-attraction, but also for complementary-attraction towards interactive computer characters (Isbister & Nass, 2000). This inconsistency of results matches the inconsistent findings regarding similarity- and complementary-attraction in human-human interaction (e.g. Franzoi, 1996).

In accordance with these findings in the context of media equation theory it can be assumed that the speaker's gender may trigger gender stereotypes and other social principles, which in turn influence the perception and evaluation of the speakers. As a result, a male respectively female speaker might be seen as an expert or a somehow likeable person and this should increase the learner's inclination to follow Grice's cooperation principle, should subsequently foster sense-making processes, and finally should result in better learning outcomes. Thus, from a *social-motivational* perspective, the speaker's gender can have several implications for learning.

On the one hand, regarding gender stereotyping, men are typically perceived as more knowledgeable, stronger, and more competent regarding technical and mathematical issues. As we used a mathematical domain in the current study, learners should be more motivated to listen to a male speaker and put more effort to sense-making activities, which in turn should improve learning outcomes. On the other hand, women are commonly seen as nicer and warmer compared to men. With respect to this latter gender stereotype, learners should be more motivated to listen to a female speaker and put more effort in trying to understand the verbal explanations, which in turn should improve learning outcomes. Thus, learning outcomes should be moderated by the particular gender stereotype that is being applied towards the speaker.

Moreover, when taking into account principles of interpersonal attraction, there may be interactions between the gender of the learner and the speaker's gender. According to the concept of similarity-attraction learners should prefer listening a speaker of the same gender and might thus show better learning outcomes when being given the opportunity to do so. However, according to the concept of complementary-attraction, learners might as well prefer listening to a dissimilar person and thus might benefit from listening to a speaker of the opposite gender. Again, learning outcomes would be moderated by the particular stereotype that is being applied towards the speaker. To conclude, there are reasonable explanations from a social-agency perspective to expect better learning outcomes for male or female speakers and for a speaker of the same or opposite gender, respectively.

On the contrary, from a *cognitive* perspective there is no obvious reason why male versus female voices should result in different learning outcomes. Both kinds of voices are equally common and understandable. The processing of a male versus a female voice requires an equal amount of cognitive resources, which implies the same amount of free cognitive resources for deeper processing. Thus, from a cognitive point of view, no effects of the speaker's gender on learning outcomes are expected.

Experiment

The aim of the presented study was twofold. First, the impact of cognitive versus social-motivational factors on learning was investigated by comparing male versus female voices. These voices differ from a social-motivational perspective, but not from a cognitive one. Second, a practical question regarding the voice principle was addressed, namely whether using a male or a female human voice for presenting narration along with animations would be more appropriate.

Method

Participants 84 students (42 female, 42 male) of the University of Tuebingen, Germany, participated in this study for either course credit or payment. Average age was 25.58 years.

Materials and Procedure A hypermedia learning environment on probability theory was used for experimentation, which consisted of four parts: a short technical instruction to the system and to the experiment, a short introduction to the domain of probability theory, a learning phase with eight worked-out examples, and a subsequent test phase. At the beginning of the experiment a questionnaire was used to measure participants' prior knowledge. In the domain introduction the basic notion of random experiments and the general rationale behind calculating probabilities were explained. In the subsequent learning phase participants had to acquire knowledge on four different problem categories, whereby each category was explained by means of two worked-out examples. The worked-out examples were presented auditory as narrations, which were accompanied by animations illustrating the respective problem statement as well as each solution step. The animations depicted objects and relations described in the problem statement in a concrete way (cf. Scheiter, Gerjets, & Catrambone, 2006). The narrated animations were learner-controlled in that learners could start, stop, and replay them. Time management was left to the learners. Depending on the experimental condition participants received a narration that was spoken by a male or a female speaker. Subsequent to the example-based learning phase the participants had to fill out several paper-based questionnaires. Finally, they had to work on a short exam with 11 test problems that were embedded in the hypermedia learning environment. The instructional materials were no longer available during problem solving.

Design and Dependent Measures The learner's gender and the speaker's gender were varied between subjects as independent variables, resulting in a 2 x 2 design.

For the variation of the speaker's gender the narrations were either spoken by male or female speakers. To counterbalance the effects of speaker-specific characteristics, three different male and three different female voices were recorded and randomly assigned to learners within the respective experimental conditions. The speed of the narrations showed only small differences among the six speakers and was synchronized with the animations. All speakers were standard-accent native speakers and their voices were adjusted in loudness.

As dependent variables social-motivational as well as cognitive variables were measured. For the evaluation of the different speakers, we used a German translation of the Speech Evaluation Instrument (SEI) by Zahn and Hopper (1982) in the short version used by Mayer et al. (2003). The SEI comprised the three subscales superiority, attractiveness, and dynamism, whereby larger values indicated higher superiority, higher dynamism, and higher attractiveness. The scales ranged from 1 to 8. Additionally, the motivation with respect to the speaker was assessed. For this purpose we designed a questionnaire called the Speaker Impression Questionnaire (SIQ) in reference to the Subject Impression Questionnaire. The latter has been developed by Deci and Ryan as a variation of the Intrinsic Motivation Inventory (Ryan, 1982) to assess the motivation towards another person also participating in an experiment. The SIQ contained six subscales: Relatedness, interest, perceived choice, pressure, effort and value. The range of the scale was from 1 to 7, whereby higher values represented higher relatedness, interest etc. Cognitive load was measured by a modified version of the NASA-TLX (Hart & Staveland, 1988) that had been successfully used as an instrument for assessing cognitive load in former studies (Gerjets, Scheiter, & Catrambone, 2004). The scale contained separate subscales for intrinsic, extraneous, and germane cognitive load. Additionally, we included a question that asked for the amount of exertion to process the voice, respectively, the understandability of the speaker. The range of the rating scale was from 0 to 10, whereby high values reflected high cognitive load and high amount of exertion to process the voice. All questionnaires were presented as paper-pencil-based versions directly after the learning phase.

Learning success was assessed by 11 test problems of varying transfer distance that were embedded in the last section of the hypermedia learning environment. For each of the test problems, one point was assigned for a correct answer; no partial credits were given. The problem-solving performance was expressed as the percentage of correct answers. Additionally, learning time spent on studying the animations (in seconds) was measured.

Moreover, several control variables were registered including socio-demographical data, prior-knowledge (measured by a multiple choice questionnaire on important concepts and definitions from the field of probability theory), and intrinsic motivation (measured by a shortened version of the Intrinsic Motivation Inventory, Ryan, 1982).

Results

The data were analyzed by means of a 2 x 2 ANOVA with the learner's gender and the speaker's gender as between-factors. For the analysis of problem-solving performance we included two additional control variables, namely the final high school grade and intrinsic motivation, resulting in a 2 x 2 ANCOVA. Both covariates showed a significant correlation with problem-solving performance (final high school grade: $r = -.36$, $p = .001$, better school grades were associated with better learning outcomes; intrinsic motivation: $r = .26$, $p = .02$, higher intrinsic motivation was associated with better learning outcomes), but were independent from each other ($r = -.01$, $p = .94$). For means see Table 1.

Effects for the Learner's Gender An analysis of the SIQ-subscales showed that male learners indicated a higher feeling of choice to listen to the speaker ($F(1, 80) = 4.15$, $MSE = 2.33$, $p = .05$), that is, male learners felt more freedom to listen to the speaker compared to female learners. Analysis of cognitive load measurements indicated that male learners rated the task difficulty (intrinsic cognitive load) as being lower in contrast to female learners ($F(1, 80) = 6.17$, $MSE = 4.63$, $p = .02$). The analysis of problem-solving performance revealed that male learners achieved better learning outcomes (cf. Figure 1) compared to female learners ($F(1, 78) = 5.86$, $MSE = 379.84$, $p = .02$). Additionally, male participants spent less time on studying the animations compared to female learners ($F(1, 80) = 4.70$, $MSE = 164223.44$, $p = .03$). No further significant effects for the learner's gender could be observed (all $Fs < 1$; except SIQ-pressure: $F(1, 80) = 1.55$, $MSE = .93$, $p = .22$, SIQ-value: $F(1, 80) = 1.45$, $MSE = 1.73$, $p = .23$, germane cognitive load: $F(1, 80) = 1.54$, $MSE = 3.35$, $p = .22$, extraneous cognitive load: $F(1, 80) = 1.38$, $MSE = 1.35$, $p = .24$).

Table 1: Means as a function of the learner's gender and the speaker's gender

	Speaker's gender			
	Male		Female	
Learner's gender	**Male**	**Female**	**Male**	**Female**
SEI:				
- superiority	5.23	5.36	5.87	5.35
- attractiveness	5.49	5.50	6.18	6.21
- dynamism	4.27	4.22	4.74	4.49
SIQ:				
- relatedness	3.80	4.08	4.54	4.15
- interest	3.81	3.78	4.71	4.24
- perceived choice	4.52	4.02	4.90	4.05
- pressure	2.12	2.29	2.07	2.43
- effort	3.56	3.00	3.67	4.21
- value	4.69	5.29	4.88	4.98
Cognitive load:				
- intrinsic	2.62	3.81	2.98	4.12
- germane	1.60	2.36	1.81	2.02
- extraneous	.67	1.05	.69	.90
Exertion to process the voice	.67	.64	.60	.90
Performance (% correct)	48.44	33.97	53.96	47.62
Learning time (in seconds)	858.19	911.48	706.76	1037.05

Effects for the Speaker's Gender The data for the speaker rating (SEI) showed a significant speaker effect for the subscale attractiveness ($F(1, 80) = 7.42$, $MSE = 1.41$, $p = .01$). Female speakers were perceived as being more attractive compared to male speakers. Analyses of the SIQ-subscales, moreover, showed that learners were more interested in listening to a female speaker ($F(1, 80) = 4.29$, $MSE = 2.28$, $p = .04$) and that they invested more effort in listening to a female speaker ($F(1, 80) = 3.99$, $MSE = 2.28$, $p = .05$). Most important, the analysis of learning outcomes revealed that learners listening to a female speaker showed a better problem-solving performance (cf., Figure 1) compared to learners listening to a male speaker ($F(1, 78) = 4.52$, $MSE = 379.84$, $p = .04$).

No further significant effects for the speaker's gender could be detected (all $Fs < 1$; except SEI-superiority: $F(1, 80) = 1.85$, $MSE = 1.12$, $p = .18$; SEI-dynamism: $F(1, 80) = 3.50$, $MSE = .83$, $p = .07$, SIQ-relatedness: $F(1, 80) = 2.69$, $MSE = 1.30$, $p = .11$).

Interactions Between Learner's and Speaker's Gender The analysis revealed no significant interactions between the learner's and the speaker's gender (all $Fs < 1$; except SEI-superiority: $F(1, 80) = 1.96$, $MSE = 1.12$, $p = .17$, SIQ-relatedness: $F(1, 80) = 1.83$, $MSE = 1.30$, $p = .18$, SIQ-effort: $F(1, 80) = 2.76$, $MSE = 2.28$, $p = .10$, learning time: $F(1, 80) = 2.45$, $MSE = 164223.44$, $p = .12$).

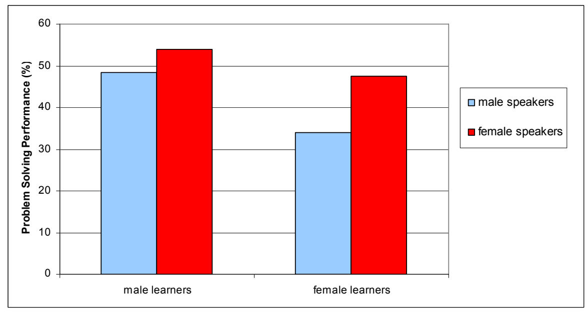

Figure 1: Problem-solving performance as a function of the learner's gender and the speaker's gender

Summary and Discussion

The presented study addresses the effect of using a male or a female voice for narrations accompanying animations depending on the learner's gender. Irrespective of the speaker's gender, the results demonstrate a superiority of male learners in that male learners required less learning time, reported less cognitive load, and performed better in the subsequent problem-solving test. These findings are in line with prior findings obtained with the hypermedia learning environment used for experimentation. They confirm also prior research on gender differences in quantitative (mathematical) abilities (for an overview see Halpern, 1992).

With respect to the speaker's gender the data revealed a bias in the speaker evaluation in favor of female speakers in that female speakers were rated as being more attractive than male speakers. Additionally, learners showed a higher motivation for listening to female speakers in terms of reporting higher interest and willingness to invest effort into studying the instructional narrations. This is in line with the gender-stereotype "women are nicer than men". Finally and most important, learners listening to female speakers showed better problem-solving performance. Thus, there is a speaker/gender effect with regard to learning outcomes that is in line with the stereotype applied.

As argued in the theoretical section, one would not expect any differences for learning from narrated animations presented by a male versus a female speaker from a purely cognitive view. Thus, the speaker/gender effect clearly indicated that a social-motivational perspective needs to be taken into account, when designing instructional multimedia messages. The pattern of findings is therefore in accordance with the assumptions of the social agency theory (Mayer et al., 2003).

Although the speaker/gender effect demonstrates the importance of social factors for learning, the relationships between cognitive, social-motivational, and performance variables are, however, not fully understood. Although our findings are in line with the five steps assumed by the social agency theory, the connection between motivation and learning outcomes remains unclear. Motivation per se does not improve learning; rather additional learning activities like activation, concentration, or sophistication of strategies need to be considered as mediators (Vollmeyer, Rollett, & Rheinberg, 1997) and should be addressed in further studies.

Another issue for further research pertains to the role of individual preferences. An alternative interpretation of the found speaker/gender effect is based on the assumption that there are individual preferences for a specific speaker. If most learners preferred a female speaker, the probability to get one's preferred speaker would be much higher in the case of a female speaker. Thus, learners may have achieved better learning outcomes, because they had been given the opportunity to listen to their individually preferred speaker and not because of stereotyping or some other social principle. In that case, individual preferences for a female speaker might have served as the mediator for the found speaker/gender effect. If this line of reasoning was correct, the speaker/gender effect should disappear if people can choose the speaker by themselves, because in this case everybody gets his or her preferred speaker - which implies that the speaker's gender should not make a difference any longer.

To sum up, the reported study demonstrates the significance of voice features for the design of narrated animations. The found speaker/gender effect provides strong support for the impact of social factors for learning. Thus, the prevailing purely cognitive approaches should be augmented by social factors. For practical design considerations, the speaker/gender effect suggests using female voices irrespectively of the learner's gender for a mathematical domain.

References

Deci, E. L., & Ryan, R. M. *Subject Impression Questionnaire.* Retrieved February 17, 2004, from www.psych.rochester.edu/SDT/measures/intrins_scl.html

Franzoi, S. L. (1996). *Social psychology.* Dubuque, IA: Brown & Benchmark Publishers.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science, 32*, 33-58.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.). *Syntax and semantics* (Vol.3). New York: Academic Press.

Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed). Hillsdale, NJ: Erlbaum.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland.

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues and user chracteristics. *Interbational Journal of Human Computer Studies, 53,* 251-267.

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*, 419-425.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*, 358-368.

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity attraction, and consistency attraction. *Journal of Experimental Psychology: Applied, 7*, 171-181.

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology, 27*, 864-876.

Reeves, B., & Nass, C. (1996). *The media equation*. Stanford, California: CSLI Publications.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450-461.

Rummer, R., Schweppe, J., Scheiter, K., & Gerjets, P. (2006). *Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts.* Manuscript submitted for publication. [Learning with multimedia: The cognitive foundations of the modality effect].

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior, 22,* 9-26.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychological Review, 10*, 251-296.

Tversky, B., Bauer Morrison, J., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human Computer Studies, 57*, 247-262.

Vollmeyer, R., Rollett, W., & Rheinberg, F. (1997). How motivation affects learning. *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 796-801). London: Erlbaum.

Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument (SEI). *Journal of Language and Social Psychology, 4*, 113-122.