International Journal of Teaching and Learning in Higher Education 2013, Volume 25, Number 1, 59-65
http://www.isetl.org/ijtlhe/ ISSN 1812-9129

How Instructional Format (Voice and Modality) Influences Subsequent Performance: An Empirical Study

Sandra Carpenter, Harry Delugach, Letha Etzkorn, and Dawn Utley
The University of Alabama in Huntsville

Educators aspire to use the most effective teaching materials possible. Their choices, ideally, should be informed by empirical evidence on how materials they use can affect students' outcomes or performance, particularly their post-instruction performance or application. The current research tested the extent to which voice (active vs. passive) and modality (oral vs. written) of instructional examples influence (a) the degree to which a new process is learned and (b) subsequent performance. Results indicate that although voice and modality may be influential during the instruction process itself, they may have minimal impact on subsequent performance. The results are discussed in the context of the importance of using empirical evidence to inform educators' choices, as well as presenting strategies for evaluating the effectiveness of learning.

Educators attempt to create instructional materials and activities that both engage and inform learners. They choose techniques that are personally appealing, embraced by learners, and popular. Popularity, however, does not necessarily imply that a technique or delivery mechanism is effective in promoting learning. As an example, in the past three decades, a great deal of focus has been placed on learning styles, which refer to individual differences in ways that diverse people learn new information and skills. A relatively comprehensive review (Coffield, Moseley, Hall, & Ecclestone, 2004) describes 71 different models of learning styles. In spite of the popularity of this concept, a recent review of the research on learning styles (Pashler, McDaniel, Rohrer, & Bjork, 2008) concludes that very few methodologically sound studies have found any support for the concept that learning styles and method of instruction influence performance. Pashler et al. (2008) state that "the contrast between the enormous popularity of the learning-styles approach within education and the lack of credible evidence for its utility is, in our opinion, striking and disturbing" (p. 117). Given limited time and resources, educators should focus their attention on variables that have been shown, through research, to improve learning. With that goal in mind, the present research addresses the question of whether the format of practice examples during instruction impacts outcomes (i.e., subsequent performance).

The use of practice examples in education and training is ubiquitous. Educators assign homework, writing exercises, and problem-solving experiences to facilitate the learning process. Some engage learners in simulations in which they are required to practice their new skills (e.g., leadership, conflict resolution, operating equipment). Two variables that might influence the effectiveness of practice examples during the learning process were investigated in the present research: (a) the modality (visual vs. auditory presentation) of practice examples and (b) the voice (active vs. passive voice) of the practice items.

Using English as the language of instruction presents special considerations. The mode (visual or auditory) and particularly voice (active or passive) used in the instruction offer unique challenges. Research shows that less educated English speakers (those who dropped out before completing high school) had more difficulty fully understanding sentences written in passive voice than active voice (Dabrowska & Street, 2006). They found that these difficulties had little connection to limitations on working memory, testing strategies, or intelligence. Their conclusion was that not all native English speakers share an innate universal grammar background.

This has further implications globally. As of 2000, there were 372 million native English speakers in the world, making English the second most common native language (Wallruff, 2000). It is the second most frequently occurring second language in the world, with 375 million non-native English speakers. When paired with the previous information, this makes English the second most influential global language (Weber, 1999). In 2007 English was considered the dominant language used in a business context, and this trend is likely to continue for the next several decades (Fishman, 1999). Today, given the increase in Internet use for communication across the world, English is the most used language of all Internet users at 27.3% (Internet World Statistics, 2010). These reasons make the present research of vital importance to anyone conducting instruction in today's environment.

In English, sentences can be constructed in either active or passive voice. For example, "Lydia calculated the statistics" is active, whereas "the statistics were calculated by Lydia" is passive. Active voice is typically used during conversation, unless special emphasis is meant to be placed on the object of the sentence (Anisfeld & Klenfort, 1973). Special emphasis might occur in the previous example if there is a question about which task was undertaken by Lydia. Passive voice is also more likely to be used in scientific or technical writing (Connatser, 2007). Research has explored whether the voice in which information is presented can affect memory. For example, Bacharach (1972) found better recall of active than passive sentences, but only for sentences that had lower semantic content (i.e., less meaning). In contrast, File and Jew (1973) found no impact of voice (passive or active) on recall. The current research will investigate whether the *voice* of the examples used during instruction affects instructional outcomes; half of the participants received the examples in the passive voice, the other half received the examples in active voice.

A second aspect of teaching materials is the sensory modality (e.g., visual, auditory, tactile) of presentation. Some research (Conway & Gathercole, 1987; Goolkasian, Foos, & Eaton, 2009; Rummer & Schweppe, 2005) shows better recall if information presentation is auditory rather than visual. Other, more classic studies, however, indicate that reading makes for better recall than listening (Dixon, Simon, Nowak, & Hultsch, 1982; McDowd & Botwinick 1984; Taub 1975). In a multimedia study, Moreno and Mayer (1999) found that students learned better when visual and verbal materials were presented in physically close proximity. Similarly, a meta-analysis of 43 independent studies was conducted by Ginns (2006), who found that a modality combination of visual graphics and spoken text resulted in the highest performance. The proposed reason for this was due to the reduced cognitive load required when using multiple modes of information processing. An Australian study, however, showed that there were certain conditions in which using auditory instruction in combination with visual input does not benefit the learning process due to limited cognitive capacity (Kalyuga, 2000). It is clear that the debate continues as to the effectiveness of modality, even combinations of modalities, on learners' performance.

Overview of the Research Problem and Methodology

The current experiment was designed to evaluate whether the *voice* and/or the *modality* examples used during instruction impact learning during instruction and subsequent performance. Participants were taught to use a method previously unknown to them—concept mapping—to represent (i.e., model/depict) a process already known to them. That is, instruction of concept mapping was used simply as the *vehicle* in our research to evaluate the potential impacts of the voice and the modality of examples used during instruction. We chose concept mapping as the vehicle for the research because it has not only content ("what"), but also process ("how"), and can be applied to any type of current knowledge (e.g., content and processes) of a learner.

For half of the participants, examples used during the learning phase of concept mapping were presented in the auditory/aural mode and for the other half of participants' examples were presented in the visual (i.e., written) mode. Half of the learning examples were presented in active voice, and the other half presented in passive voice. The overall design was therefore a 2 (voice: active vs. passive) x 2 (modality: aural vs. written) between-participants factorial.

The experiment was composed of two phases: the instruction phase and the application phase. In the instruction phase, participants learned to use concept mapping to model processes already known to participants (e.g., grocery shopping). Participants were assigned to one of the four instructional conditions, based on how the information to be modeled was presented (i.e., voice and modality conditions). In the application phase, the degree to which they had learned the rules of conceptual mapping was assessed. Participants were asked to create a concept map of "printing a syllabus using WebCT" (a course management/delivery system). Rather than giving the participants information about this process, participants were expected to create a concept map based on their prior knowledge of using WebCT. The quality of participants' models in the application phase was expected to be the combination of (a) how well participants had learned the concept mapping rules during the instruction phase, and (b) their prior knowledge of using WebCT. A more detailed description of the participants, materials, and procedure is provided in the next section.

Method

Participants were 78 undergraduate students enrolled in introductory psychology courses at The University of Alabama in Huntsville. These courses are general education options for students from all majors, providing a wide diversity of student participants. The sample included 55 women and 23 men, had a mean age of 21 years (SD = 4.6), and was composed of 60.3 % White, 19.2 % Black, 2.6 % Hispanic, 5.1 % Asian, and 12.8% individuals of other ethnic heritages. English was the first language of all participants.

The instructional materials included written descriptions about how to use "sticky notes" to create a concept map of a sequence of steps in a shopping experience. The goal of the research was to determine whether participants showed any differences in learning concept mapping as a function of the voice and/or modality of the examples used during learning. All participants reported that they had never used concept mapping previously. Thus, the participants' task was to learn to use concept mapping to represent information that they already knew.

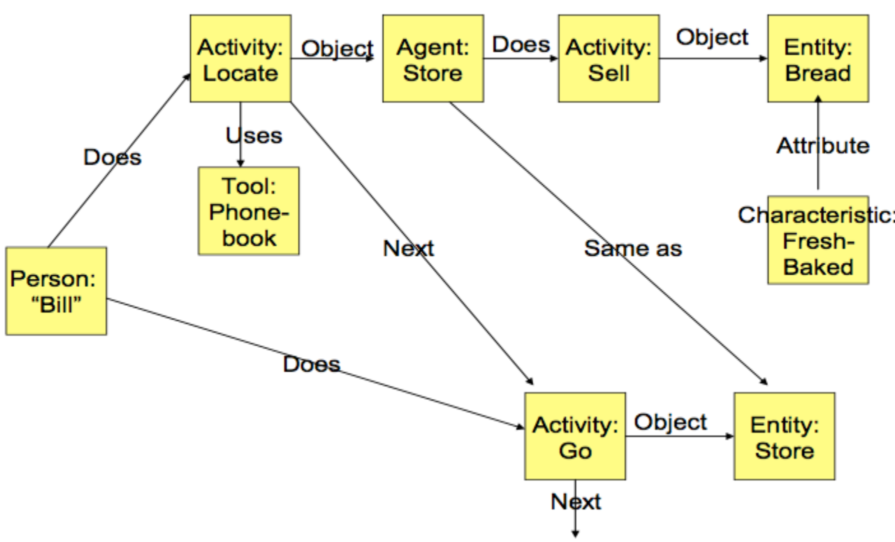
The concepts that participants could use were: entity (a noun; e.g., "banana"), tool (a type of entity used in an activity; e.g., "knife"), agent (an active entity; e.g., "monkey"), person (a particular type of agent), characteristic (a property of something; e.g., "yellow"), process (a complex action; e.g., "select"), and activity (a single, simple action; e.g., "peeling a banana"). The relations participants could use included: attribute, object, recipient, uses, is a, output, input, next, same as, or, part of, and does. One of the several examples given at the beginning of instruction was "Bill eats the banana." The concept map of this example is shown in Figure 1. Other examples showed relations between entity and attributes, tools and activities, etc. Each participant was given a set of the instruction materials to use throughout the training and experimental sessions. They could write notes on the materials if they wished. The relation between agents or entities and activities was mapped in the same way, regardless of whether the sentence to be mapped was in the active or passive voice. Thus, "Bill eats the banana" and "the banana is eaten by Bill" would be mapped in the same way (Figure 1).

Participants engaged in three instruction session stages related to the process of grocery shopping. During each stage, they were given two sentences to model using sticky-note concept maps. Half of the participants were given written sentences, which they retained during the experimental session. The other half of the participants were read the sentences out loud by the researcher. The length of the sentences was not expected to overtax memory, given their reading grade level and brevity. Half of the participants were given the description of the process in the active voice, the other half were presented the passive voice. Thus, there were four experimental +conditions (2 x 2). As examples of the active voice conditions, some sentences were: in the active voice, "Bill uses a phone book to locate a store that sells fresh-baked bread" and "Then he gives money to the cashier." In the passive voice condition, analogous sentences were: "To locate a store that sells fresh-baked bread, a phone book is used by Bill" and "Then the cashier is given money."

Expert concept maps, showing the concepts and relations that should be present in all participants' concept maps, were developed by the experimenters to be used as feedback for the participants. A concept map used as feedback for the first stage of instruction is shown in Figure 2. After the first and second stages of instruction, participants were shown the expert concept maps to compare to their own and were given the opportunity to ask questions. This feedback was designed to increase learners' understanding of the concept mapping rules. Participants retained these

Figure 1
Example of a Concept Map in the Instruction Phase

Figure 2
The Expert Concept Map Used as feedback for the First Phase of Instruction



expert maps to use as aids in creating subsequent maps. The quality of the concept maps created by participants in the third stage of instruction was taken as evidence of the degree to which they had learned the rules for creating concept maps. To assess the quality of their third concept maps, these maps were compared to the researchers' expert concept map of the third stage process. The entire instruction session, with feedback, took approximately 30 minutes.

Following the instruction phase, participants engaged in the application phase. During this phase they were asked to use concept mapping to model a process with which they all had experience: printing a syllabus using WebCT. Unlike the instruction phase, participants were not given detailed sentences to use during modeling; rather, they needed to generate their own cognitive representation of the process and translate it into concept maps again using sticky-notes and the rules learned during the instruction phase.

After the application phase, participants were given a quiz designed to evaluate the degree to which they had accurate and comprehensive knowledge of the process for which they had been asked to create a concept map: printing a syllabus using WebCT. Although all participants were expected to have printed a syllabus using WebCT, they were expected to have varying degrees of experience with this process. The quiz consisted of four multiple-choice questions about the structure of WebCT (e.g., "Which courses could be accessed through this course management system?") and a set of four items to be rank ordered in terms of the sequence needed to print a syllabus (e.g., "click the syllabus icon," "choose the appropriate course"). The quiz questions thus assessed whether students knew and remembered the steps in printing a syllabus and whether they recalled the correct sequence of the steps, which was the same knowledge they had been asked to represent in their concept maps. Scores on the quiz could range from 0-8. Finally, participants were asked to report whether they had ever printed a syllabus using WebCT. As expected, all participants reported that they had. Thus, participants during the application phase were attempting to use their new knowledge of concept mapping to represent their previous knowledge of using WebCT.

Results

Two coders compared the scores they derived from students' concept maps of the instruction (e.g., grocery shopping) and application (e.g., WebCT) phases with an expert concept map created by the researchers (e.g., see Figure 2 for an expert concept map). In the experts' three instructional maps (one for each of the three phases of instruction), 13 concepts, 20 relations, and 13 linked concepts (defined below) were possible for participants' instruction phase concept maps. The correlation between coders (averaged across the three measures) was $r = 0.91$ for the third instruction map and $r = 0.84$ for the WebCT concept maps, indicating good inter-rater reliability.

To test the hypothesis that voice and modality would influence the quality of the concept maps from the third instruction stage, the number of concepts and relations from participants' concept map of the third stage were used as dependent variables. In some cases, participants placed a concept in their concept map, but did not connect it to another concept through a relation. A third variable, linked concepts, was therefore created and was based on the number of paired concepts that were appropriately linked by the appropriate relation (according to the expert concept map). An example of this can be seen in Figure 1, with the concept of "monkey" linked to the concept of "eat" by the relation "does." Each variable was submitted to a 2 (voice: passive vs. active) x 2 (modality: written vs. oral) between subjects ANOVA.

For the number of concepts, there were no significant effects. For the number of relations there was one reliable effect, a significant interaction between voice and modality, $F(1, 74) = 5.23$, $p < 0.03$. Follow-up analyses indicated that there was no difference in the number of relations appearing in participants' concept maps in the active voice condition, but that there were significant differences for the passive voice condition, $t(1,37) = 2.46$, $p < 0.02$, with participants in the written modality generating the "best" concept maps.

For the number of appropriately linked concepts, again there was only the one significant interaction, $F(1,74) = 5.04$, $p < 0.03$. Follow-up analyses indicated that there was no difference in the number of appropriately linked concepts appearing in participants' concept maps in the active voice condition, but that there were significant differences for the passive voice condition $t(1, 37) = 2.46$, $p < 0.02$, with participants in the written modality again generating the "best" concept maps.

To test the hypothesis that voice and modality would influence the quality of the concept map in the application phase, the number of concepts, relations, and linked concepts from participants' concept maps of printing a syllabus were used as dependent variables. None of the main effects or interactions was statistically significant for any of the three variables.

A regression analysis was performed to evaluate the extent to which the concept maps of the WebCT process, printing a syllabus, were the hypothesized combination of the participants' (a) understanding of the concept modeling process and (b) prior knowledge of WebCT. For each participant, the number of concepts from the third instruction phase and the score on the WebCT knowledge quiz were used as predictors of the number of concepts generated for the WebCT concept maps. Both the instructional concept map (β = 0.43, $p = 0.001$) and the quiz (β = 0.22, $p < 0.03$) were significant predictors of the number of concepts accurately appearing in the WebCT concept map. In addition, the number of linked concepts was significantly predicted by the accuracy of the third instruction concept map (β = 0.46, $p < 0.001$) and the quiz (β = 0.28, $p < 0.005$). Means and standard deviations for each dependent variable, for each condition, are shown in Table 1.

Discussion

It is clear that both voice and modality affected the quality of participants' concept maps during the instruction phase. Participants who were given the process to model through written sentences in the passive voice had better models than participants in the oral presentation in the passive voice condition. We speculate that the active voice is easier to comprehend and remember in either written or auditory situations, as we hear people speaking to us much more frequently than we read. People are less likely to have experience listening to sentences in the passive voice than reading in the passive voice (e.g., in technically written documents). Possibly, then, participants had some difficulty retaining the information that needed to be modeled even though the sentences were simple and brief. Also, active sentence structure may be more engaging to listeners, whereas passive structure may sound technical and stilted. Educators and trainers may therefore want to choose active voice for their instruction if that is an alternative for them.

This pattern, however, did not generalize to the application phase, where participants used their own cognitive representation of a process, then modeled it. Thus, although voice and modality of practice examples affected instruction performance somewhat, they had no impact on later application of the information learned during instruction. The finding that voice and modality did not affect post-instruction performance outcomes is heartening. The important, take-home lesson for educators is they likely do not need to spend undue time or attention on determining how these variables might impact the effectiveness of their instruction. Rather, their resources would be better targeted at other educational practices that have been shown, through empirical research, to influence post-instruction performance.

A limitation to this experiment is that, prior to the instruction phase, one example showing an agent, Bill, was presented in the active voice to all participants (as shown in Figure 1). This may have impacted learning somewhat for participants in the passive voice conditions. It is unlikely to have had a substantive effect, however, as it was only one of the types of maps used during instruction. Moreover, participants in the passive voice/written condition had overall better instruction maps, making it unlikely that poorer performance in the passive voice, oral condition was solely due to one of the examples being in the active voice.

This paper evaluates the effectiveness of instruction by having learners engage in a post-instruction task relevant to the training. The effectiveness of instruction is often conducted during or at the end of instruction sessions, rather than being evaluated in terms of post-instruction performance in the work environment (Tyler, 2002). Reactions (e.g., satisfaction) of trainees to the trainer's presentation (usually conducted through questionnaires) is the most frequently occurring evaluation method, used approximately 78% of the time, with assessment of actual learning conducted only 32% of the time. Transfer of instruction—application of the lessons learned during instruction to the job site—is rarely measured (i.e., only 9% of the time), although this is the long-term goal of the instruction process. Therefore,

Table 1
Mean Number of Concepts and Relations Appearing in Concept Maps as a Function of Voice and Modality

	Active Voice		Passive Voice	
Modality	*M*	*SD*	*M*	*SD*
Number of Relations				
Oral	7.30	2.72	6.10	3.01
Written	6.33	3.34	8.20	2.70
Number of Linked Concepts				
Oral	5.80	3.12	4.85	2.70
Written	4.56	3.35	6.55	2.35
Number of Concepts (ns)				
Oral	10.55	2.89	9.70	2.72
Written	10.33	2.38	11.40	2.56

educators are encouraged to evaluate the effectiveness of instruction procedures that they might consider using in terms of their impact on subsequent applications of the instruction. In the present research, voice and modality affected performance on the practice examples, but not the post-instruction measure of learning. Thus, evaluating outcomes of instruction may not be reflective or predictive of the application of the learning. The goal of the present research was to add to the evidentiary base of research on aspects of instruction that influence post-instruction outcomes-to identify whether voice or modality have systematic impacts on learning. The tentative conclusion is that they do not, but replications should be conducted to verify this pattern for different topics being learned under various learning conditions. Moreover, other modalities should also be considered. As indicated by Jewitt Kress, Ogborn, and Tsatsarellis (2001), learning is accomplished through the interaction of linguistic, visual, and action modes, with our research focusing primarily on the linguistic component of learning.

References

Anisfeld, M., & Klenbort, I. (1973). On the functions of structural paraphrase: The view from the passive voice. *Psychological Bulletin*, *79*(2), 117-126. doi:10.1037/h0033970

Bacharach, V. R. (1972). Semantic and syntactic constraints on free-recall learning of sentential material. *Journal of Experimental Psychology*, *96*, 223-225. doi:10.1037/h0033476

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review*. London, UK: Learning and Research Skills Centre.

Connatser, B. R. (2007). Two centuries of progress in technical communication. *Journal of Technical Writing and Communication*, *37*(2), 129-149. doi:10.2190/Y2T7-4672-8U65-U85T

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*(3), 341-361. doi:10.1016/0749-596X(87)90118-5

Dabrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*(6), 604-615. doi:10.1016/j.langsci.2005.11.014

Dixon, R. A., Simon, E. W., Nowak, C. A., & Hultsch, D. F. (1982). Text recall in adulthood as a function of level of information, input modality, and delay interval. *Journal of Gerontology*, *37*(3), 358-364. doi:10.1093/geronj/37.3.358

File, S. E., & Jew, A. (1973). Syntax and the recall of instructions in a realistic situation. *British Journal of Psychology*, *64,* 65-70. doi:10.1111/j.2044-8295.1973.tb01327.x

Fishman, J. A. (1998-1999). The new linguistic order. *Foreign Policy*, *113*, 22-26+34-40. doi:10.2307/1149230

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511-525. doi:10.1016/j.learninstruc.2006.10.001

Goolkasian, P., Foos, P., & Eaton, M. (2009). Modality effects in sentence recall. *Journal of General Psychology*, *136*(2), 205-223. doi:10.3200/GENP.136.2.205-224

Internet World Statistics. (2010*). Internet world users by language: Top 10 languages*. Retrieved from http://www.internetworldstats.com/stats7.htm

Jewitt, C., Kress, G., Ogborn, J., & Tsatsarellis, C. (2001). Exploring learning through visual, actional, and linguistic communication: The multimodal environment of a science classroom. *Educational Review, 53*(1), 5-18. doi:10.1080/00131910120033600

Kalyuga, S. (2000). When using sound with a text or picture is not beneficial for learning. *Australian Journal of Educational Technology*, *16*(2), 161-172.

McDowd, J., & Botwinick, J. (1984). Rote and gist memory in relation to type of information, sensory mode, and age. *Journal of Genetic Psychology*, *145*(2), 167-178. doi:10.1080/00221325.1984.10532264

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*, 358-368. doi:10.1037/0022-0663.91.2.358

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*, 105-119. doi:10.1111/j.1539-6053.2009.01038.x

Rummer, R., & Schweppe, J. (2005). Evidence for a modality effect in sentence retention. *Psychonomic Bulletin and Review*, *12*(6), 1094-1099. doi:10.3758/BF03206449

Taub, H. A. (1975). Mode of presentation, age, and short-term memory. *Journal of Gerontology*, *30,* 56-59. doi:10.1093/geronj/30.1.56

Tyler, K. (2002). Evaluating evaluations. *Training and Development*, *47*, 85-93.

Wallruff, B. (2000, November). What global language? *The Atlantic Monthly*, *286,* 52-66.

Weber, G. (1999). The world's 10 most influential languages. *National Bulletin*, *24*, 22-28.

SANDRA CARPENTER, PhD, is a Professor of Psychology at The University of Alabama in Huntsville. She obtained her PhD in Social and Personality Psychology from the University of

California, Santa Barbara. She engages in interdisciplinary research on cognitive aspects of teamwork, training, and cybersecurity.

HARRY DELUGACH, PhD, is an Associate Professor of Computer Science and Director of the Honors Program at The University of Alabama in Huntsville. He received his MS in computer science from the University of Tennessee, Knoxville and his PhD in computer science from the University of Virginia. His research interests include knowledge modeling, logic, database semantics, software engineering, and team mental models.

LETHA ETZKORN, PhD, is a Professor of Computer Science at The University of Alabama in Huntsville. She has a BEE and an MSEE (Electrical Engineering) from the Georgia institute of Technology and an MS and PhD in Computer Science from UAH from the University of Alabama in Huntsville. In addition to studying mental models in teamwork, her primary research areas are in software engineering and artificial intelligence applications to software engineering.

DAWN UTLEY, PhD, is an associate professor in the department of Industrial Systems Engineering and Engineering Management at The University of Alabama in Huntsville. She obtained her MS in Industrial Systems Engineering and Engineering Management at the University of Tennessee in Knoxville and her PhD in Industrial Systems Engineering and Engineering Management at The University of Alabama in Huntsville. Her research interests include engineering management, teamwork, motivation of the technical workforce, and organizational metrics.