To cite, please refer to: Chevalier, J. W. & Fox Tree, J. E. (2012). Using hetero-modal communication to optimize knowledge and awareness. *American Journal of Psychology, 125*(3), 361-368.

# Using Hetero-Modal Communication to Optimize Knowledge and Awareness

Jessica W. Chevalier and Jean E. Fox Tree

Department of Psychology, UC Santa Cruz

Author Note

This research was supported by faculty research funds granted by the University of California Santa Cruz. We thank Robert W. Proctor and two anonymous reviewers for comments on an earlier draft of this manuscript. Correspondence can be addressed to Jean E. Fox Tree, foxtree@ucsc.edu, Psychology Department, Social Sciences II room 277, University of California, Santa Cruz, CA, 95064.

Abstract

Accurate, error-free communication is essential for success in many areas, such as eyewitness testimony, human factors design, business, education, and personal relationships. Traditional communication employs similar modalities: participants communicate by talking, or by writing, but not both at the same time with the same addressees. New communicative technologies have broadened this vista. For example, one communicator can speak and the other can type. We tested communicative effectiveness using accuracy and error-detection in a trivia recall test, evaluating the roles of presentation and retrieval modalities on reporting facts stored in long-term memory. Hetero-modal communication (hearing and writing or reading and saying) was more effective than homo-modal communication (hearing and saying or reading and writing), with the most correct responses and the most errors caught. This has direct connections to communicative success and applied tests of skill.

Using Hetero-Modal Communication to Optimize Knowledge and Awareness

Historically, communication took place in similar modalities. People spoke face-to-face, voice-to-voice (telephone), letter-to-letter, or Morse code-to-Morse code. With modern technology, hetero-modal communication has become increasingly common. A single setting can combine live meetings, telephone and conference calls, web cameras, instant messaging, and email. We tested some possible costs and benefits of these new means of communication, focusing on the distinction between what we label *homo-modal* forms, hearing and saying or reading and writing, versus what we label *hetero-modal* forms, hearing and writing or reading and saying. Specifically, we tested whether the homo-modal versus hetero-modal settings differentially affected people's accuracy at responding to questions about factual knowledge or their abilities to catch their own mistakes. Catching mistakes was tested with Moses illusion questions (Erickson & Mattson, 1981; Hannon & Daneman, 2001): Even though it is common knowledge that it was Noah, not Moses, who took animals on the ark, participants often respond *two* to the question *How many animals of each kind did Moses take on the ark?*

**The Importance of Optimizing Knowledge and Awareness**

Knowing how to communicate accurately and how to effectively catch errors is important in many arenas. For example, accuracy and error detection are critical for eyewitness testimonies. Eyewitnesses are notorious for not remembering correctly what happened, yet their testimony is often instrumental in determining whether or not someone will go to jail; criminal cases without forensic support can still see conviction rates of 90% with the presence of a single eye witness (Haber & Haber, 2000). These testimonies are given in a single modality, either written or verbal question-and-answer. Given the importance of eyewitness testimony, it is worth asking whether this is the optimal method or if a hetero-modal format might be more veridical.

Optimizing accuracy and error detection is also critical to human factors design. GPS devices are increasingly popular and often designed with the potential to be used homo- or hetero-modally. Many devices allow the driver to speak commands and hear directions or type commands and read directions (homo-modal functionality). However, they often also offer the potential to speak commands and read directions or type commands and hear directions (hetero-modal). While it is important to know where you are and where you are going, for safety reasons the most important focus of a driver's attention should be the road and other vehicles and potential obstacles on it. Therefore communication with a GPS device should be as accurate and minimally demanding of drivers' attentional resources as possible.

Many other products and behaviors also rely on optimizing accuracy and error detection. For example, human error potential in nuclear power plant control rooms (which can lead to such disasters as Three Mile Island or Chernobyl) can be decreased by factors of 2 to 10 if workstations incorporate standard human engineering concepts (Swain & Guttman, 1983). As another example, the ability to catch and correct errors is essential for education. If learners do not know when they are wrong, they cannot build on their knowledge in order to be right more often.

**Multiple Resource Model**

Carefully controlled human factors studies of non-communicative behavior, such as laboratory tracking in flight simulation, consistently found hetero-modal advantages (see Wickens, 1984 for a comprehensive list; see also Wickens, Sandry, & Vidulich, 1983; Wickens, 2002). These results are predicted by Wickens' (2002) *multiple resource model* which proposes that variance in time-sharing performance is accounted for by four categorical and dichotomous dimensions, of which perceptual modalities is one dimension with multiple levels (including auditory and visual perception). According to the model, if the resource demand or task difficulty is equal for two tasks that require one level of any given dimension, they will interfere with each other and performance will suffer more than when two tasks require separate levels of that dimension. Therefore, all else being equal, two tasks of equal resource demand that both require visual perception, or both require auditory perception (homo-modal), should result in poorer performance than two tasks of equal resource demand which require both visual and auditory perception (hetero-modal).

However, Wickens (2002) also pointed out that the two competing visual channels can (1) require scanning between them if they are far apart or (2) impose confusion and masking if they are close together, and that either of these can result in degraded performance. This possibility suggests that superior hetero-modal performance, rather than being the result of separate perceptual resources in the brain, could instead reflect a homo-modal disadvantage. But there is other evidence that supports a hetero-modal advantage rather than a homo-modal disadvantage. Subvocalization is more common hetero-modally than homo-modally (Locke & Fehr, 1972). Subvocalization is the recoding of stimuli into phonetic features, a process which makes stimuli more resistant to decay or interference. This suggests improved performance in hetero-modal tasks. The tasks we employ involve neither scanning between far-apart items nor confusion from close-together items. Thus, in our study, superior performance in the hetero-modal conditions would be the result of a hetero-modal advantage rather than a homo-modal disadvantage.

**Error Monitoring**

In contrast to both the hetero-modal advantage and homo-modal disadvantage proposals, there is reason to expect error monitoring to only be improved when saying a response out loud, whether in the hetero-modal read-say setting or the homo-modal hear-say setting. The reason for this prediction is that speaking out loud involves both proposed production monitoring systems, the *internal monitor* and the *external monitor* (Levelt, 1983, 1989). The internal monitor checks pre-articulatory speech for errors before anything is uttered. The external monitor detects errors in post-articulatory speech, and is activated through hearing one's own speech. Thus, the opportunities for correction increase once material has been both produced and heard. Nonetheless, because error correction with either monitor can occur quite rapidly (Motley, Camden, & Baars, 1982), we predict that the larger differences observable in the current study will be with hetero-modal versus homo-modal communication.

Although a great deal of research has been done on the internal monitor (Morgan & Wheeldon, 2003; Motley, Camden, & Baars, 1982; Wheeldon & Morgan, 2002), less is known about the external monitor. What has been observed with the external monitor is that there were more errors in production than perception and that monitoring one's own speech requires more resources than detecting errors in the speech of others (e.g. Hartsuiker & Kolk, 2001; Oomen & Postma, 2002). The external monitor can be tricked into failing to notice errors given semantic similarities between the correct answer and an incorrect lure, as demonstrated by the Moses illusion (Erickson & Mattson, 1981; Hannon & Daneman, 2001), which generalizes over a wide range of materials and conditions (for overview see Hannon & Daneman, 2001).

**Current Study**

The current study exploited the Moses illusion to test whether accuracy and error-catching abilities are related to communicative setting, specifically whether the communication is hetero-modal (read-say or hear-write) or homo-modal (read-write and hear-say). Said another way, the study evaluated the role that presentation and retrieval modality played on reporting facts stored in long-term memory, an activity that has direct connection to both communicative success and applied tests of skill.

In the current study, people answered trivia questions based on information stored in long-term memory. Some of the questions were valid and some were anomalous. The valid questions tested common knowledge facts that most people would be reasonably expected to know even if they had not spent much time thinking about them, such as what the yellow part of an egg is called. The anomalous questions aimed to create Moses illusions. The framing of the anomalous questions included semantically similar but incorrect information that was intended to lure some participants into answering the question incorrectly. As a question was presented, participants accessed long-term memory to recognize and check for the sense of the question. Next, participants mapped the accessed memory onto a response. Either the question was interpreted as valid and they had an answer or it was interpreted as anomalous and identified as such.

This study differs from prior studies in that participants were explicitly encouraged to change their answers if they thought the initial response was incorrect. This modification enabled evaluation not only of accuracy but also of error correction. We predicted that participants would have (1) the most correct answers and (2) catch and correct the most errors when presentation and retrieval modes were hetero-modal rather than homo-modal.

This study further differs from prior studies by testing individual differences in competency, and by testing individual differences using a within-subjects design. Other studies of *unskilled and unaware* participants (Ehrlinger et al., 2008; Kruger & Dunning, 1999) used between-subjects designs, with one group of participants considered skilled and aware and another group considered unskilled and unaware. In our study, each participant’s responses were broken down into three categories of skill and awareness. Responses that were correct and unchanged were categorized as skilled and aware. Responses that were incorrect and unchanged were categorized as unskilled and unaware. No answer and *don't know* responses were categorized as unskilled and aware, as were initially incorrect responses that were changed to correct responses. In addition to considering each participant's skill and awareness levels as identified by their answers, we also assessed the relationship between a participant's overall competency, as measured by self-assessed ratings of knowledge, and their responses (correct unchanged, incorrect unchanged, no answer and *don't know*, incorrect to correct). This allowed testing of whether the benefit of hetero-modal communication was greater for individuals with more or less knowledge.

## Method

### Participants

The participants were eighty undergraduate students at the University of California at Santa Cruz who participated to fulfill a course requirement. Data for five participants were discarded because they were not native English speakers, for four participants due to their inability to follow directions (i.e. spoke answers in sections where the instructions indicated to write answers), and for two participants due to apparent inability to understand the questions, as measured by their *no answer/don't know* responses being more than two standard deviations above the mean. This left sixty-nine participants.

### Materials and Design

The study used a 2 x 2 within-treatments design. Participants answered 15 anomalous and 15 valid questions for each input type (aural or read) and each output type (spoken or written), for a total of 120 questions presented in blocks counter-balanced by input type. Each participant experienced each of the four experimental conditions in counter-balanced order: hear-say, hear-write, read-say, and read-write. Each stimulus question was also presented counter-balanced in each of the two input conditions: aural and visual. This controlled for the possibility that certain questions could be more difficult to answer in particular input or output modes.

**Apparatus**

The experiment was run on a desktop PC running Windows 2000 Professional. Stimuli were presented one by one using SuperLab 4.0.7. A Panasonic PVGS-150 video camera was used to record the experimental sessions.

**Procedure**

Participants answered 120 questions in blocks of four. Text questions were presented for a total of 8 seconds each, including time to both read and respond. Participants' responses were recorded in pen on paper given to each participant for that purpose. Participants had 4 seconds to respond to audio questions after the question finished playing. Oral responses were later transcribed from video recordings.

Participants were instructed to answer aloud or to write responses on paper as quickly as they could. They were told that they could change their answers, as long as they did so before the next question was presented. They were further informed that a one-word answer would usually be sufficient, and that if they did not know the answer or could not recall it right away, to answer *don't know*. Finally, participants were told that some questions had things wrong with them, such as in the question *How many animals of each kind did Moses take on the ark?* If they encountered a question like this, they were to answer *wrong*.

After participants finished answering questions, they rated their knowledge from 1-5 of the areas covered by the stimuli questions (e.g. history, games, pop culture). These rating were used to calculate a *sum knowledge score* and to evaluate participants' overall knowledge.

**Results**

Correct responses included both the factual responses to valid trivia questions (e.g., *What country borders the U.S. to the North?* Answer: *Canada*) and identification of anomalous questions as anomalous (e.g., *Who stabbed President Lincoln at Ford's Theater?* Answer: *wrong*). Incorrect responses included anything other than the correct response (e.g., anything other than *Canada* for *What country borders the U.S. to the North?* and anything other than *wrong* for *Who stabbed President Lincoln at Ford's Theater*?). Likewise, both correct-to-incorrect responses and incorrect-to-correct responses could be produced with both valid and anomalous questions (e.g., the correct-to-incorrect answers "Canada – no, wrong" and "wrong – wait, John Wilkes Booth" and the incorrect-to-correct answers "wrong – oh! Canada" or "John Wilkes Booth, no – wrong").

Four 2 x 2 anovas, input (read or hear) x output (write or say), were conducted to evaluate hetero-modal (hear-write and read-say) versus homo-modal performance (read-write and hear-say) for each dependent variable: (1) the mean number of initially incorrect answers changed to correct, (2) the mean number of *don't know* answers plus unanswered questions, (3) the mean number of correct unchanged answers, and (4) the mean number of incorrect unchanged answers. Subsequently we conducted four planned pooled *t*-test comparisons between hetero-modal and homo-modal conditions.

There were significant interactions between input and output modalities for three of the four tests, as follows: (1) $F(1, 68) = 9.62$, $p = .003$, for incorrect changed to correct (see Figure

1), (2) $F(1, 68) = 17.51$, $p < .001$, for *don't know* answers plus unanswered questions (see Figure 2), and (3) $F(1, 68) = 4.87$, $p = .03$, for correct unchanged (see Figure 3). There was no interaction for the incorrect unchanged responses, $F(1, 68) = 1.71$, $p = .20$. Among the four ANOVAs, there were only two main effects. Participants produced more *no answer/don't know* responses when speaking as compared to writing, 4.34 to 3.63, $F(1, 68) = 10.94$, $p = .002$, and they produced more correct unchanged responses when writing as compared to speaking, 18.59 to 17.71, $F(1, 68) = 7.48$, $p = .008$.

There were significant planned pooled *t*-test comparisons for three of the four tests, as follows: (1) More incorrect responses were changed to correct hetero-modally, $t(68) = 3.10$, $p = .003$; responses of this type indicate that a participant was able to monitor for and change an error in a timely fashion (unskilled and aware), (2) There were fewer *no answer/don't know* responses hetero-modally, $t(68) = -4.18$, $p < .001$; responses of this type indicate that a participant did not know the correct answer, but was aware of that fact (unskilled and aware), and (3) There were more correct responses hetero-modally, $t(68) = 2.21$, $p = .031$; responses of this type indicate that a participant knew the correct answer and was aware of that fact (skilled and aware). There were about the same number of incorrect responses hetero-modally as homo-modally, $t(68) = -1.31$, $p = .195$. Responses of this type indicate that the participant did not know the correct answer and lacked awareness of that fact (unskilled and unaware).

Competency as measured by sum knowledge scores was unrelated to the rate of production of (1) incorrect to correct answers, (2) correct unchanged answers, and (3) incorrect unchanged answers across either hetero-modal or homo-modal conditions (see Table 1 for test statistics). Across both hetero-modal and homo-modal conditions, sum knowledge scores were negatively correlated with *no answer/don't know* answers, revealing the unsurprising outcome that the more participants knew about the world the fewer *no answer/don't know* responses they produced (see Table 1 for test statistics).

## Discussion

In a test of answers to trivia questions, hetero-modal input and output modalities increased accuracy and error-detection compared to homo-modal communication. There were more correct responses and more errors caught and corrected hetero-modally. There were also fewer responses left blank or answered with *don't know* hetero-modally. This demonstrates that participants felt more confident with their knowledge hetero-modally.

In addition, participants produced more correct unchanged responses when writing as compared to speaking. One explanation is that participants put extra weight on their written responses, increasing their motivation to ensure those responses were correct. Another explanation is that writing takes longer, allowing participants more time to come up with the right answer. A future study could equate the time allotted for responses, perhaps by recruiting only participants who can type quickly and requiring participants to type responses. A problem with allowing participants longer time to say responses, as another way of equating response time, is that asking participants to wait several seconds before responding violates conversational norms for question-answering (Jefferson, 1989; Sacks, Schegloff, & Jefferson, 1974; Smith & Clark, 1993). As expected based on these norms, participants produced more *no answer/don't know* responses when speaking as compared to writing. One explanation for this is that participants felt pressure to respond in a timely manner, so if the answer was not readily available, they opted to say *don't know* sooner than they may have written *don't know*. Another explanation is that the clock ran out more often with spoken responses, resulting in more *no answer*s.

One take-home message from the writing main effect might be that to increase chances of being accurate, written responses should be encouraged. However, if the goal is not only to have the most correct answers but to catch the most errors as well, hetero-modal communication is better.

More generally, the current study demonstrates that in situations where respondents are aware of their knowledge, hetero-modal communication is beneficial, increasing the number of correct unchanged responses (skilled and aware), increasing the number of incorrect responses that were changed to correct responses (unskilled and aware), and decreasing the number of no answer and *don't know* responses (unskilled and aware). Only when respondents are unaware of their knowledge or mistakes does hetero-modal communication fail to help, as with the incorrect unchanged responses (unskilled and unaware). That is, while communicating hetero-modally will not give people insight they do not possess, it will optimize their ability to correctly communicate what they do know, enabling them to have greater skill and awareness.

With respect to Wickens' (2002) multiple resource model, the current study supports a genuine hetero-modal advantage, rather than a homo-modal disadvantage. With respect to the external speech monitor, we did not find evidence that error monitoring was only improved when saying a response out loud. As anticipated based on the speed of error correction (Motley, Camden, & Baars, 1982), larger differences in error-correction were observed between communication modalities (the hetero-modal advantage) rather than output modalities.

Knowing what form of communication is most effective for accurately conveying information and for catching errors is increasingly important as communicative systems extend beyond traditional homo-modal boundaries. The information obtained through the current study has direct applications for educational settings where factual information is routinely taught and tested. The communication of factual information and the detection of errors in that information could be enhanced with hetero-modal teaching and testing techniques.

The study also contributes to understanding successful communication in other domains, as multi-media formats become more and more common in business and interpersonal settings. This study tested only communication of common facts and the ability to detect anomalies, so the benefit for other contexts is still speculation at this time. Future researchers could test if hetero-modal communication is also the most effective in specialized situations such as discussing emotional topics, conducting business deals, questioning eyewitnesses, interrogations, piloting airplanes, and other human factors applications. For example, the ability to communicate accurately and detect errors is important for human interaction with GPS devices for driving directions. The current study suggests that having both an auditory and visual component should lead to more successful navigation than either alone.

References

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*(1), 98-121.

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20, 540-551.

Haber, R. N., & Haber, L. (2000). Experiencing, remembering and reporting events. *Psychology, Public Policy, and Law, 6*(4), 1057-1097.

Hannon, B., & Daneman, M. (2001). Susceptibility to semantic illusions: An individual-differences perspective. *Memory & Cognition*, 29 (3), 449-461

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, 42, 113-157.

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a standard maximum silence of approximately one second in conversation. In D. Roger & P. Bull (Eds.), *Conversation* (pp. 166–196). Clevedon: Multilingual Matters.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Locke, J. L., & Fehr, F. S. (1972). Subvocalization of heard or seen words prior to spoken or written recall. *American Journal of Psychology, 85*(1), 63-68.

Morgan, J. L., & Wheeldon, L. R. (2003). Syllable monitoring in internally and externally generated English words. *Journal of Psycholinguistic Research*, 32 (3).

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour*, 21, 578-594.

Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes*, 17 (2), 163-184.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 696–735.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, 32, 25–38.

Swain, A. D. & Guttman, H. E. (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications. Final Report. US Nuclear Regulatory Commission, Washington, DC, USA, Report No NUREG/CR-1278, 1983, 707 pp. (March 01, 1985). *Applied Ergonomics, 16*, 1, 68-68.

Wheeldon, L. R., & Morgan, J. L. (2002). Phoneme monitoring in internal and external speech. *Language and Cognitive Processes*, 17 (5), 503-535.

Wickens, C. D. (1984). Processing resources in attention, in R. Parasuraman and R. Davies (eds), *Varieties of Attention* (New York: Academic Press), 63-101.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wickens, C. D., Sandry, D. & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, output, and central processing. *Human Factors,* 25, 227-248.

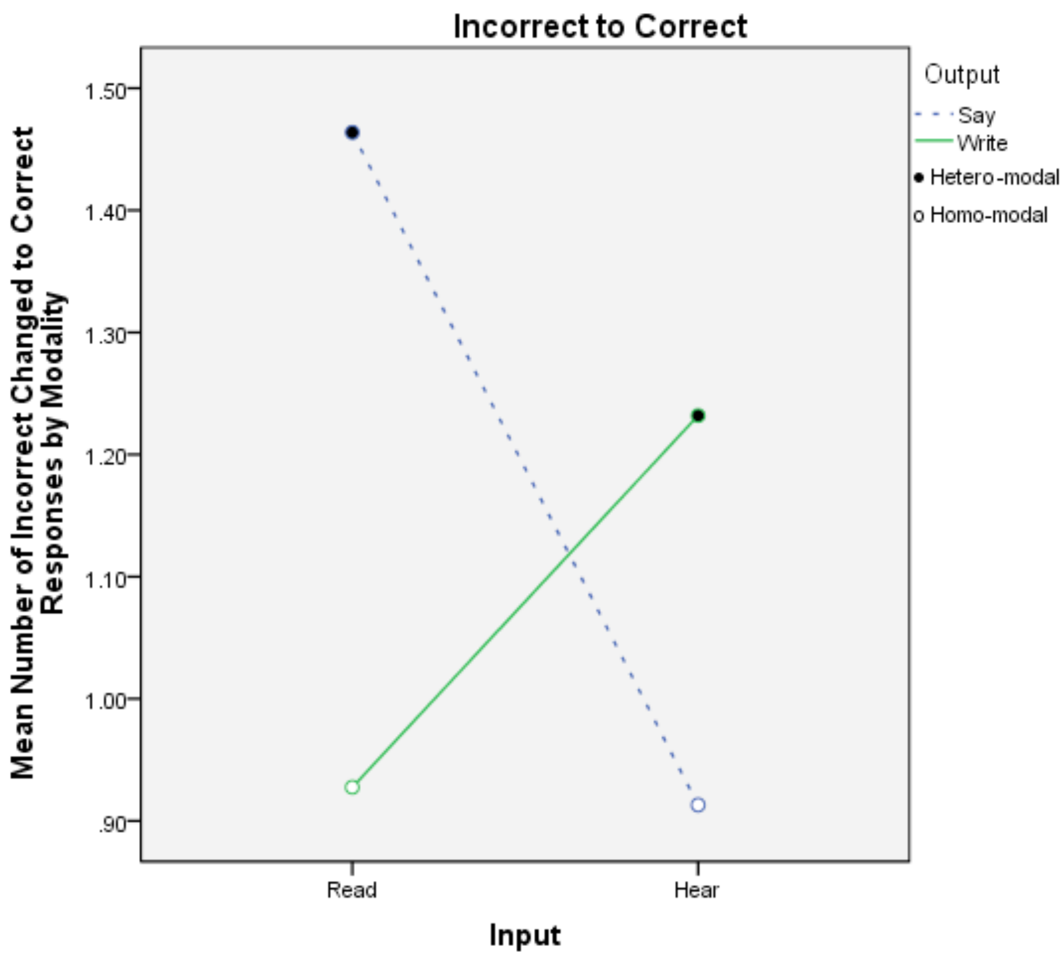


Figure 1. *Graph of incorrect to correct responses by modality*.

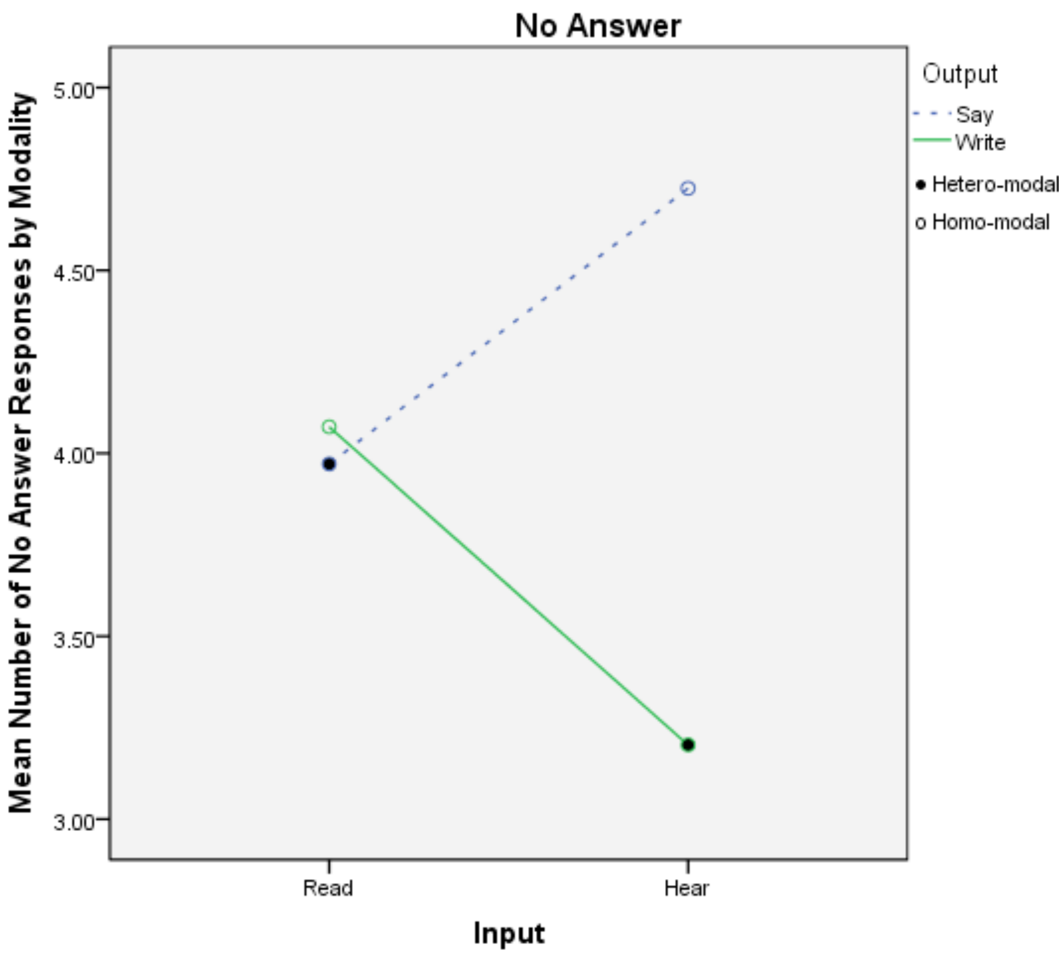


Figure 2. *Graph of no answer responses by modality*.

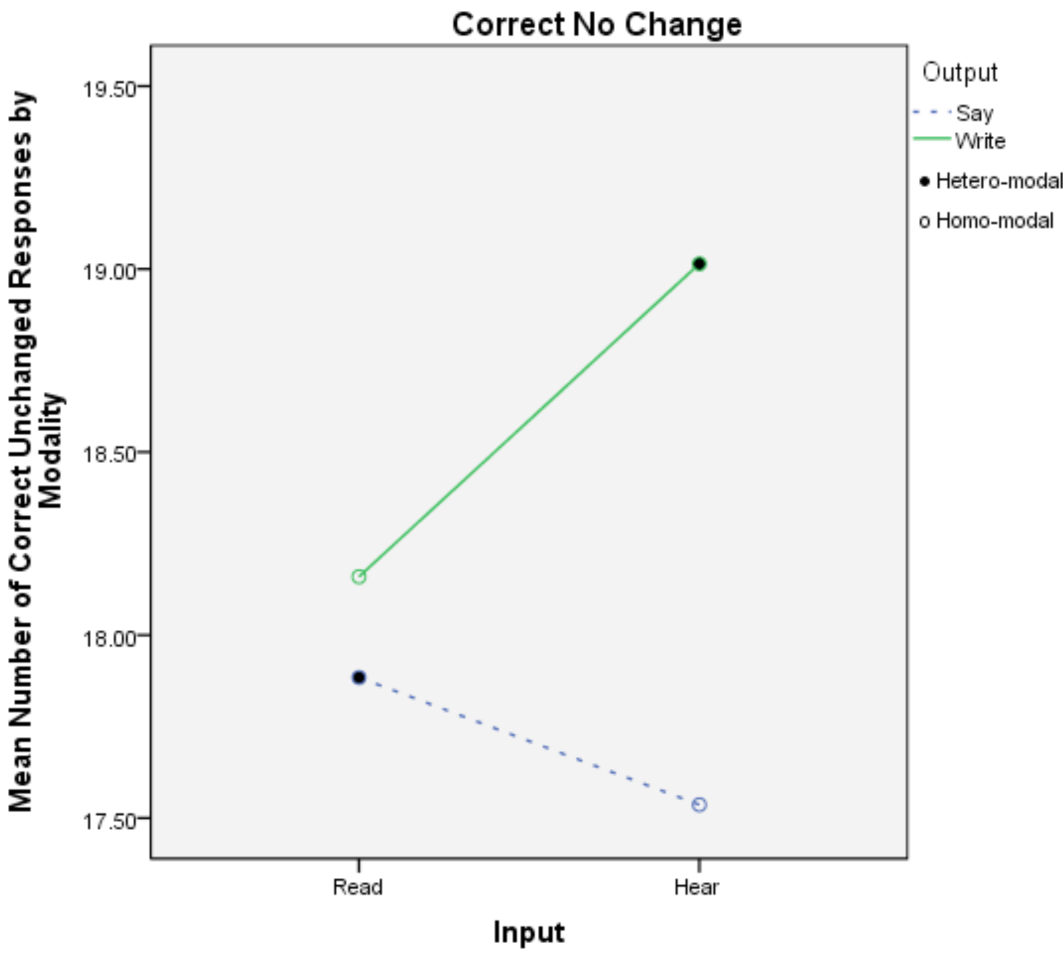


Figure 3. *Graph of correct unchanged responses by modality*.

Table 1

*Correlations for sum knowledge scores and responses by modality*

| Hetero-Modal | *r* | *p* | Homo-Modal | *r* | *p* |
|---|---|---|---|---|---|
| Incorrect to Correct | -.10 | .40 | Incorrect to Correct | .21 | .09 |
| No Answer/Don't Know | -.45 | .00* | No Answer/Don't Know | -.45 | .00* |
| Correct Unchanged | .19 | .12 | Correct Unchanged | .09 | .47 |
| Incorrect Unchanged | .04 | .75 | Incorrect Unchanged | .18 | .13 |

*Correlation is significant at the .05 level. (2-tailed).