Gestures in Communication through Line Graphs

Cengiz Acartürk (ACARTURK@Metu.Edu.Tr)

Özge Alaçam (OZGE@Metu.Edu.Tr)

Cognitive Science, Informatics Institute
Middle East Technical University, 06800, Ankara, Turkey

Abstract

Line graphs are widely used in communication settings, for conveying information about states and processes that unfold in time. The communication is achieved by the contribution of other modalities than graphs, such as language and gestures. In a set of experimental investigations, we analyzed the production and comprehension of gestures during communication through line graphs. The findings reveal a systematic use of gestures as well as the limitations of cognitive resources due to the split of attention between the modalities.

Keywords: Gesture production; gesture comprehension; graph comprehension; line graphs.

Line Graphs in Time Domain

Line graphs represent statistical data, most often the relationship between two domain variables. In line graphs, line segments are used for representing the mapping between the values. When used in time domain, line graphs represent the mapping between the values of the domain variable and time. From the perspective of human comprehension, line graphs in time domain have a peculiar characteristic: they represent not only statistical data but also *states* and *processes* that unfold in time, by providing perceptual cues for continuation (Figure 1).

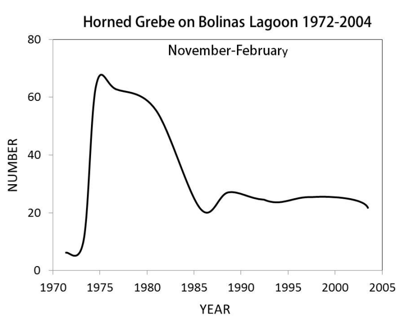

Figure 1: Sample population graph from PRBO (2012); redrawn based on the original.

Accordingly, the population graph in Figure 1 does not only represent the mapping between years and the population of the bird species but also leads to a conceptualization of how population *increases*, *decreases* or *remains stable* in certain periods of time.[1]

Human conceptualization through statistical graphs has been a topic of interdisciplinary research since the past 30 years. The research on graph comprehension has covered a broad range of analyses including the investigations on perceptual processes of graph comprehension (e.g., Cleveland & McGill, 1985), analysis from the perspective of psychology and usability (e.g., Kosslyn, 1989), cognitive models (Lohse, 1993; Peebles & Cheng, 2002), educational psychology and instructional design (Winn, 1987; Mautone & Mayer, 2007). On the other hand, the research on modalities that accompany graphs, such as language and gesture in communication through graphs, has been scarce except for a few studies (e.g., Gerofsky, 2011, on gestures in graphs of polynomial functions). Concerning the relationship between language and gestures, gestures have been considered as having a key role in organizing, conveying spatial information, and preventing decay in visuospatial working memory (Hostetter & Alibali, 2010), thus having the potential to promote learning in educational contexts (Goldin-Meadow, 2010). Analyzing the relationship between graphical cues, language and gestures, the present study investigates communication through line graphs from the perspective of multimodal interaction.

Communication through Line Graphs

Graphs are abundant both in spoken communication settings (e.g., classroom settings) and in written communication settings (e.g., newspaper articles). Communication through graphs is achieved by means of the contribution of several modalities: language (both in written form and in spoken form), graphical cues in written communication settings, and gestures in spoken communication settings. The previous research on multimodal comprehension reveals a frequent use of spatial terms that convey spatial information in communication through line graphs (Habel & Acartürk, 2007). Moreover, in spoken communication, people tend to produce more gestures when they perform tasks that involve spatial information, compared to tasks with no spatial information (Alibali et al., 2001; Trafton et al., 2006; Hostetter & Sullivan, 2011). Consequently, in communication through line graphs, humans frequently produce gestures that accompany spoken language.

---

[1] Line graphs are generated based on a set of assumptions that specify the way the data points are represented by lines. For instance, according to the original source (PRBO, 2012), the line graph in Figure 1 was generated by applying a local regression method called Loess smoothing on data points. The resulting spatial aspects of line graphs, such as smoothness, influence humans' interpretation of the states and processes (Acartürk et al., 2008), a topic beyond the scope of the present study.

Gestures in communication are of different types: the most commonly used ones are deictic (or pointing) gestures and iconic (or representational) gestures. Deictic gestures show objects, people and places, whereas iconic gestures are representations of shape of an object or an action (Özçalışkan & Goldin-Meadow, 2005). In communication settings, deictic gestures facilitate achieving joint attention on objects, whereas iconic gestures overlap with spatial tasks (Alibali et al., 2001; Trafton, et al., 2006). In communication through line graphs, humans may produce both deictic gestures and iconic gestures. It is also not unusual that humans emphasize certain aspects of processes and states represented by line graphs, such as a specific increase, a peak or a stable period of the domain value, in addition to emphasizing an overall pattern. Graphical annotations (also called graphical cues) on graph lines are generally used for this purpose.

The major focus of the present study is to investigate gestures in communication through line graphs, both from a production perspective and a comprehension perspective. For a systematic analysis, we limited the domain of investigation to the relationship between gestures and graphical cues in line graphs (rather than the overall pattern of the graph line). In the first step of the analysis, one group of participants produced gestures during a verbal description task (Experiment 1). We considered the produced gestures as human interpretations of the structural aspects of the states and processes represented by the graphs. The gestures produced by the participants of Experiment 1 were used for designing the stimuli for a comprehension experiment (Experiment 2). This approach resembles what has been termed the “3Ps (Preference-Production-Performance) program” as an empirical method for selecting appropriate representations for abstractions (Kessell & Tversky, 2011). The two approaches are similar; in that, both aim to perform an empirical investigation of the representations rather than leaving the decision for selecting the appropriate representation to intuitions of the graphic designer. Instead of graphic representations, however, we investigated gestures in communication through graphs in a set of consecutive analyses (i.e., the outcome of Experiment 1 was used for preparing the stimuli set in Experiment 2).

Experiment 1

In Experiment 1, the participants presented verbal descriptions of annotated graphs. Spontaneous gestures of the participants were analyzed in terms of the relationship between the type of the graphical cue and the gesture type.

Participants, Materials and Design

A total of seven participants (*Mean age* = 25.4, *SD* = 3.78) who were graduate students or teaching assistants from the Faculty of Education, Middle East Technical University (METU) participated in the experiment, five of which reported having teaching experience. The experiment language was Turkish, which was the native language of the participants. The participants were asked to imagine themselves in an online meeting, in which their task was to present single-sentence summaries of annotated graphs to the audience. According to the scenario, the audience was able to see the participant (i.e., the presenter) but not the graphs. Therefore, the presenter first investigated the graph displayed on a computer screen, then s/he turned towards the audience (an audience picture displayed on another computer screen), and then presented a single-sentence summary of the graph. The participants were not informed that their gestures were in the focus of the experiment. Each participant presented the single-sentence summaries for 14 annotated graphs, thus generating 14 video recordings per participant. The graphs represented populations of bird species in a lagoon. Each graph involved a graphical annotation that emphasized a certain aspect of the information represented, such as a specific increase or a peak. In particular, three types of annotations were used.

- Process annotation: A diagonal arrow that emphasized a specific increase or a decrease.
- Durative state annotation: A horizontal arrow that emphasized a specific period of constant value.
- Punctual state annotation: A point-like circle that emphasized a specific value such as a peak value.

The 14 stimuli involved 2 graphs for familiarization of the participant to the task. The remaining 12 stimuli involved 6 punctual state annotations (2 for the start point of the lines, 2 for middle and 2 for the endpoint of the lines), 4 (diagonal) process annotations and 2 (horizontal) durative state annotations (Figure 2).

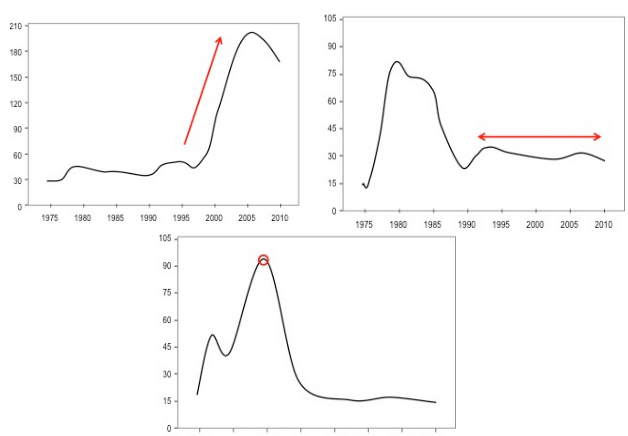

Figure 2: Sample annotated graphs with a process annotation (upper left), a durative state annotation (upper right), and a punctual state annotation (bottom).

Following Gerofsky (2011), we employed the coding scheme proposed by Creswell (2007) for the analysis of (14 graphs x 7 participants) 98 experiment protocols. The Noldus Observer XT event logging software was used for coding. Two coders analyzed the protocols according to the following criteria: For each gesture in the video recording, the coder first classified the gesture in terms of its directionality: having no gesture, no direction, being vertical, horizontal, diagonal or other.[2] Then the coder identified the following features of each gesture: size (small or big), palm direction (up, down or front), speed (slow or fast) and start position (low, middle or high). In the present study, we focus on the directionality of gestures by leaving the analysis of other features to an extended study. One coder initially coded the entire data, and a second coder, who was blind to the hypothesis, carried out 57% of the dataset. Interrater reliability between coders was calculated by Cohen's kappa. The results revealed an agreement value of .78. According to Landis and Koch (1977), a value above .61 indicates substantial interrater agreement.

Results

The participants gestured in 86% of the protocols. This number is close to what Hegarty et al. (2005) reported: the participants gestured when they described solutions to mental animation problems in 90% of the cases.[3] Pearson's chi square test and follow-up McNemar tests were conducted to investigate the relationship between the annotation type and the gestures produced by the participants. The test showed a significant effect of annotation type on gesture, $\chi^2 = 48.1$, $p < .05$. In particular, for the graphs with process annotations, the participants produced more vertical and diagonal gestures compared to both horizontal gestures, $\chi^2 = 15.7$, $p < .05$, and gestures with no direction, $\chi^2 = 5.88$, $p < .05$. On the other hand, they produced more horizontal gestures compared to other types of gestures for durative states, $\chi^2 = 4.08$, $p < .05$. Finally, for punctual annotations, more non-directed pointing gestures were produced compared to vertical gestures, $\chi^2 = 16.5$, $p < .05$, to horizontal gestures, $\chi^2 = 26.0$, $p < .05$, and to diagonal gestures, $\chi^2 = 20.8$, $p < .05$.

These findings show that, in terms of the categorization of the gestures (cf. McNeill, 2005; Özçalışkan & Goldin-Meadow, 2005) the participants produced iconic gestures for process annotations and durative state annotations. On the other hand, for punctual state annotations, they produced pointing gestures that were ambiguous between iconic (because the pointing gesture was representational) and deictic (by definition).

Experiment 2

The findings obtained in Experiment 1 suggest that humans produce a specific type of gesture depending on the emphasized aspect of the information represented in the graph. Based on the results obtained in Experiment 1, we investigated comprehension of gestures by humans in Experiment 2. For this, we prepared 14 video recordings in which a narrator presented a single-sentence summary of annotated graphs by producing a relevant gesture concurrently with the spoken description. The verbal description was a single-sentence summary for a graph with process annotation, a graph with durative state annotation or a graph with punctual state annotation. For example, for a graph with a process annotation, the narrator uttered the sentence "[t]he population of coot in the lagoon increased between 1980 and 1985" while producing an upward diagonal gesture that showed an increase. She uttered the sentence "[t]he sanderling population in the lagoon remained stable between 1975 and 1985" accompanied by a horizontal gesture for a graph with a durative state annotation. Finally, for a graph with a punctual state annotation, the narrator uttered the sentence "[t]here exists about 120 terns in the lagoon in the year 2010" accompanied by a pointing gesture (Figure 3). The duration of the video recordings was between 5.3 seconds and 8.6 seconds ($M = 6.24$, $SD = 0.95$).

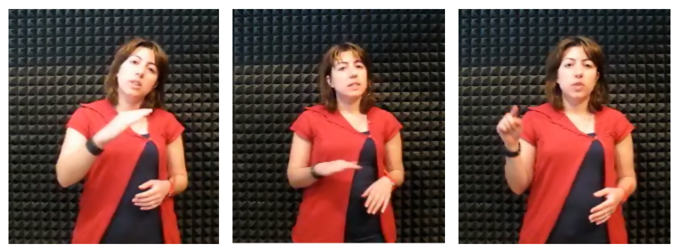

Figure 3: Snapshots from the video recordings with a diagonal gesture for a process annotation (left), a horizontal gesture for a durative state annotation (middle), a pointing gesture for a punctual state annotation (right).

Experiment 2 was conducted in three different conditions. In the first condition, the participants played the videos on the screen one by one and they listened to a single-sentence summary for each graph concurrently. In the second condition, the participants played the same video recordings but the sound was muted, therefore they interpreted what was presented on the screen only. In both the first condition and the second condition, we noted that participants' gaze shifted between the gesture and the face of the narrator. We interpreted this finding as a potential source of attention split. Therefore, in the third condition, we provided the participants with only gestures not the face of the narrator. In all conditions, the participants were asked to predict the described graph among a set of three alternative graphs.

Condition 1: Concurrent Interpretation of Gestures and Language

Participants, Material and Design. Eleven participants (*Mean age* = 31.8, *SD* =5.1), who were either graduate or undergraduate students of METU, participated in the experiment. Each participant was presented 14 video recordings (2 trials and 12 tests). After playing each recording, the participant was asked to choose the described

[2] The 'other' category involved beat gestures (simple up-and-down movements without semantic information) or more complex gestures like the combination of vertical, horizontal or diagonal movements.

[3] A further investigation revealed that the five participants who reported teaching experience gestured in 93% of the protocols whereas the two participants who reported no experience in teaching gestured in 68% of the protocols. The finding suggests a potential correlation between teaching experience and gesturing.

graph among three alternatives (the alternate graphs were the same except for the graphical annotation). After submitting each choice, the participant reported a subjective evaluation for confidence ("*How confident are you about your judgment?*") by using a *1* to *3* scale (*1* showing a low confidence, *3* showing a high confidence; Beattie and Shovelton, 1999). The stimuli were displayed on a Tobii non-intrusive 120 Hz eye tracker, integrated into a 17" TFT monitor with a resolution of 1024x768 pixels. The spatial resolution and the accuracy of the eye tracker were 0.25° and 0.50° respectively. No time limit was set for the answers. The order of presentation of the stimuli was randomized.

Results. The participants exhibited high success rates in predicting the annotated graphs, for all three types of gestures, i.e. the process gesture (M = 1.0, i.e. 100%), the durative state gesture (M = 1.0) and the punctual state gesture (M = .93, SD = 0.01). The results of an ANOVA test revealed a significant difference between the gesture types, $F(2, 20) = 5.17$, $\eta^2 = .36$, $p < .05$: the success rate in punctual states was lower than the other two gesture types. A comparison of the confidence scores reported by the participants, however, revealed no significant difference between the gesture types $F(2, 20) = 1.86$, $\eta^2 = .16$, $p > .05$. Finally, the participants spent the longest time to answer punctual state questions (M = 7.03 seconds, SD = 2.97), which was longer than both processes (M = 5.27 seconds, SD = 1.69) and durative states (M = 6.65 seconds, SD = 2.97), $F(2, 20) = 3.61$, $\eta^2 = .27$, $p < .05$, without a significant difference between the last two.

Condition 2: Interpretation of Gestures

The first condition of the experimental investigation employed the most naturalistic setting for an online communication environment: the participants listened to the narrator when she produced the gestures concurrently. In other words, both modalities (i.e., language and gesture) were available to the participants. Therefore, it is not possible to analyze the role of language and gestures separately in comprehension of the presented stimuli. The participants might have used the linguistic information to predict the graph without taking the gestures into account. In the second condition of the study, we asked the participants to predict the described graphs by displaying the video recordings with the sound muted.

Participants, Material and Design. Eighteen participants, from METU participated in the experiment (*Mean age* = 21.1, SD = 1.37). They were presented the same video recordings but they did not hear the narrator. The same experimental procedure was applied as in the previous condition.

Results. The participants in Condition 2 exhibited high success rates for processes (M = .93, SD = .11) and durative states (M = .91, SD =. 19) but a significantly lower success rate for punctual states (M = .55, SD = .22), $F(2, 34) = 25.4$, $\eta^2 = .60$, $p < .05$. The difference between processes and durative states was not significant. The lack of the language modality resulted in significant differences between the three gesture types in confidence scores, $F(2, 34) = 18.1$, $\eta^2 = .51$, $p < .05$. The participants reported lower confidence scores for punctual states (M = 2.01, SD = 0.42) compared to both processes (M = 2.61, SD = 0.33) and durative states (M = 2.61, SD = 0.47). As in Condition 1, the mean response time of the participants in punctual states (M = 4.34 seconds, SD = 1.74) was longer than both processes (M = 2.66 seconds, SD = 0.80) and durative states (M = 2.88 seconds, SD = 1.44), $F(2, 34) = 10.5$, $\eta^2 = .38$, $p < .05$, without a significant difference between the last two.

Condition 3: Attention Split between Gestures and Face

The findings obtained in Condition 1 and Condition 2 show that the lack of linguistic information results in lower success rates in predicting the answers; in particular, in punctual states. The analysis of the eye movements of participants revealed another finding about inspection patterns on the video recordings: the participants shifted their gaze between narrator's gestures and face both in Condition 1 (M = 2.55, SD = 0.28) and in Condition 2 (M = 2.68, SD = 0.35), without a significant difference between the two groups of participants, $F(1, 26) = 0.83$, $p > .05$, thus suggesting a potential source of attention split during comprehension. Therefore, a third group of participants were presented narrator's gestures only, without face and sound.

Participants, Material and Design. Twenty-one participants (*Mean age* = 21.2, SD = 2.37) from METU participated in the experiment. The participants were presented the same stimuli except that the video recordings were cut from the top, so that only the gestures (but not the face) of the narrator were displayed. The same experimental procedure was applied as in the previous conditions.

Results. The participants showed high success rates for processes (M = .96, SD = .10) and durative states (M = 1.0) but a relatively lower success rate for punctual states (M = .70, SD = .19), $F(2, 40) = 32.0$, $\eta^2 = .61$, $p < .05$, without a significant difference between processes and durative states. Confidence scores for punctual states (M = 2.31, SD = 0.40) were also significantly lower than both processes (M = 2.69, SD = 0.30) and durative states (M = 2.76, SD = 0.37), $F(2, 40) = 14.7$, $\eta^2 = .42$, $p < .05$. Finally, they spent the longest time to answer punctual state questions (M = 3.52 seconds, SD = 1.18), significantly different than both processes (M = 2.60 seconds, SD = 1.04) and durative states (M = 2.41 seconds, SD = 1.02), $F(2, 40) = 8.84$, $\eta^2 = .31$, $p < .05$, without a significant difference between the last two.

A comparison between the three groups of participants in the three conditions of Experiment 2 showed that the highest success rate (in predicting the correct annotated graph that was described in the video recording) was obtained when the participants listened to the single-sentence description of the graphs while playing the video recording. The lack of the language modality, however, resulted in a decrease in success rates. On the other hand, helping the participants to focus on gestures only (by removing narrator's face from the view) resulted in an increase in the success rates (Figure 4).

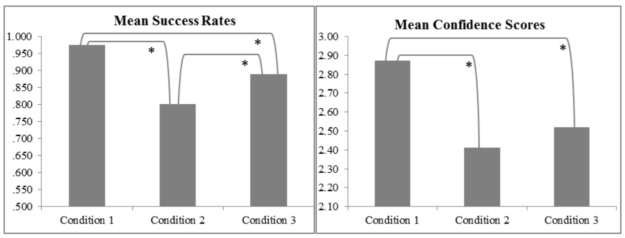

Figure 4: Mean success rates (left) and mean confidence scores (right) in Experiment 2.

For the comparison of the results obtained in the three conditions of Experiment 2, a Games-Howell test was applied since the number of samples for the three groups was not equal and the population variances were significantly different. The test returned a significant difference between the three groups of participants in their overall success rates, $F(2, 47) = 17.2$, $\eta^2=.42$, $p < .05$. Finally, a comparison of the confidence scores between the participant groups showed that the lack of the language modality resulted in lower self-confidence of the participants about their predictions, $F(2, 47) =10.3$, $\eta^2=.30$, $p < .05$ (Figure 4).

Discussion

In two experiments, we investigated how humans produce gestures (Experiment 1) and comprehend gestures (Experiment 2) when they communicate through graphically annotated line graphs. In Experiment 1, the participants produced more frequent vertical and diagonal gestures to emphasize processes (e.g., *increase*, *decrease*) whereas they produced more horizontal gestures to emphasize durative states (e.g., *remain stable*). Those two types of gesture are known as *iconic gestures* and they overlap with representation of spatial information (Alibali et al., 2001; Trafton et al., 2006). For emphasizing punctual states (e.g., a *peak*), the participants produced *pointing gestures*. In Experiment 2, three groups of participants were presented video recordings and they were asked to predict the described graphs: the video recordings were designed based on the correspondence between diagonal gestures and processes, between horizontal gestures and durative states, and between pointing gestures and punctual states. When gestures were displayed concurrently with linguistic information (Condition 1), the participants showed a high success rate in all gesture types. When language modality was absent, however, they showed a lower success rate and lower self-confidence, in particular in punctual states. These findings suggest a low efficiency of the pointing gesture (in the form of a deictic pointing gesture) in conveying information about punctual states. On the other hand, vertical and diagonal gestures were efficient in conveying information about processes. Horizontal gestures were efficient in conveying information about durative states. An explanation to these findings may be related to the major roles of iconic gestures and deictic gestures in communication. In contrast to iconic gestures that convey spatial information, the major role of pointing gestures is to attract the attention of the communication partner (McNeill, 2005; Özçalışkan & Goldin-Meadow, 2005). Consequently, further research is needed to identify more appropriate candidates for emphasizing punctual states in graphs. For instance, a circular movement of the index finger might be more appropriate for representing punctual states.

Another finding obtained in Experiment 2 was that participants' back and forth movement of their gazes between the gestures and the face of the narrator is a potential source of attention split during the course of comprehension. Although speech sound was absent (Condition 2) and therefore no linguistically useful information was provided by the narrator (except for the possibility of lip reading), the participants shifted their gaze several times between the narrator's face and the gestures. When the narrator's face was removed from the video recordings (Condition 3), an increase in success rates was observed compared to Condition 2, though the success rates were still lower than the ones obtained when the linguistic information was available (Condition 1). Although this is far from being a naturalistic setting for communication through graphs, the analysis of such boundary cases is necessary for understanding the contribution of separate factors in comprehension. In fact, the findings support the likelihood of the split of attention. A possible explanation may be sought in the domain of the intersection between cognitive science and instructional science, in which the previous studies show that the split of attention between information sources leads to degraded learning outcomes due to limited cognitive resources that are available for understanding the instructional material (Sweller et al., 1998; Mayer & Moreno, 1998). Consequently, the findings suggest that tasks demands may be high in communication through graphs; therefore, attention split should be avoided by, for instance, using small window sizes so that the communication partner is able to attend to both gestures and face in a single fixation.

Conclusion and Future Work

In communication settings, humans produce gestures when they convey spatial information. As a consequence, in communication through line graphs, gestures are an indispensable part of communication. In this study we investigated how humans produce and comprehend gestures in communication through line graphs. We found that vertical and diagonal gestures efficiently convey information about processes such as increase and decrease, and horizontal gestures efficiently convey information about durative states. However, pointing gestures are not efficient in conveying information about punctual states, possibly due to their concurrent role as deictic gestures in communication. Our future research will address finding more appropriate gesture candidates for punctual states. The future research will also address the investigation of the interaction between gestures and gradable (scalar) adjectives, gradable adverbs and spatial prepositional phrases and adverbials, e.g. *from*, *to*, and *between*, which are part of the vocabulary in communication through line graphs, in addition to state verbs and verbs of change.

Acknowledgments. We thank METU HCI Research and Application Laboratory for their technical support. We also thank four anonymous reviewers for their helpful comments and suggestions. We thank Christopher Habel for stimulating discussions on graphical annotations.

References

Acartürk, C., Habel, C., & Çağıltay, K. (2008). Multimodal comprehension of graphics with textual annotations: The role of graphical means relating annotations and graph lines. In J. Howse, J. Lee & G. Stapleton (Eds.), *Diagrammatic Representation and Inference: LNCS* (Vol. 5223, pp. 335-343). Berlin/Heidelberg: Springer.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169-188.

Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, 18, 438-462.

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, 229, 828-833.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches (2nd ed.)*. Sage Publications.

Gerofsky, S. (2011). Mathematical learning and gesture: Character viewpoint and observer viewpoint in students' gestured graphs of functions. *Gesture, 10*(2-3), 321-343.

Goldin-Meadow, S. (2010). When gesture does and does not promote learning. *Language and Cognition, 2*(1), 1-19.

Habel, C., & Acartürk, C. (2007). On reciprocal improvement in multimodal generation: Co-reference by text and information graphics. In I. van der Sluis, M. Theune, E. Reiter & E. Krahmer (Eds.), *Proceedings of the workshop on multimodal output generation* (pp. 69-80). University of Aberdeen, UK.

Hegarty, M., Mayer, S., Kriz, S., Keehner, M. (2005). The role of gestures in mental animation. *Spatial Cognition and Computation*, 5, 333-356.

Hostetter A. B., Alibali M. W. (2010). Language, gesture, action! A test of the gesture as simulated action framework. *Journal of Memory and Language*, 63, 245-257.

Hostetter, A. B., & Sullivan, E. L. (2011). Gesture production during spatial tasks: Its not all about difficulty. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. Austin, TX.

Kessell, A. M. & Tversky, B. (2011). Visualizing space, time, and agents: Production, performance, and preference. *Cognitive Processing*, 12, 43-52.

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3*(3), 185-226.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human-Computer Interaction, 8*(4), 353-388.

Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology, 99*(3), 640-652.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.

McNeill, D. (2005). Gesture and thought. University of Chicago Press.

Özçalışkan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, 96, B101-B113.

Peebles, D. J., & Cheng, P. C.-H. (2002). Extending task analytic models of graph-based reasoning: A cognitive model of problem solving with Cartesian graphs in ACT-R/PM. *Cognitive Systems Research*, 3, 77-86.

PRBO (2012). *Waterbird Census at Bolinas Lagoon, Marin County, CA*. Public report by Wetlands Ecology Division, Point Reyes Bird Observatory (PRBO) Conservation Science. http://www.prbo.org/cms/366, retrieved on January 29, 2012.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial Cognition and Computation, 6*(1), 1-29.

Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration* (Vol. 1, pp. 152-198). New York: Springer-Verlag.