Learning from a Scientific Discourse through Video Lessons: When Gestures can't help [1]

Francesco Ianì[1] Alessandro Lombardo[2] and Monica Bucciarelli[1-3]

[1] Dipartimento di Psicologia, Università di Torino, Torino, Italy
[2] Choralia, Milano, Italy
[3] Centro di Logica, Linguaggio, e Cognizione, Università di Torino, Torino, Italy
francesco.iani@unito.it

Abstract. The classical literature on co-speech gestures has proved their beneficial role in comprehension and learning from discourse. However, while the role of gestures in narrative discourse comprehension has been widely explored, their role in scientific discourse comprehension has been neglected and most of the literature on learning science is concerned with gestures accompanying single scientific concepts. Since instruction done by video can exploit the power of gestures, our twofold aim was to explore the effect of gestures accompanying a scientific discourse delivered through video. In three experiments we ascertained whether learning from scientific discourse through videos benefits from gestures, observed or produced. The results have revealed that comprehension and learning from a scientific discourse do not improve when the teacher gestures compared to when the teacher does not gesture (Experiments 1 and 2) and that learner's gestures, as compared to teacher's gestures, can worsen comprehension and learning (Experiments 2 and 3). These results have implications for technology enhanced learning.

Keywords: Learning, Scientific Discourse, Video Lessons, Gestures.

1 Introduction

Speakers' co-speech gestures favor deep discourse comprehension in listeners (Cutica & Bucciarelli, 2015). Many researchers have argued that the successful comprehension of a discourse is tantamount to the construction of a coherent mental model (Zwaan & Radvansky, 1998), and according to different theoretical frameworks, such representations are referred to as "situation model" (van Dijk & Kintsch, 1983) or "mental model" (Johnson-Laird, 1983). Following the tenets of the mental model theory, Bucciarelli (2007) and Cutica and Bucciarelli (2008; 2011) advanced a mental model account for the cognitive change produced by gestures: gestures, whether observed or produced, favor the construction of a mental model of the discourse they accompany. Since mental models are discrete representations in nature, the information conveyed by co-speech gestures, also represented in a non-discrete format, can be easily incorporated into the discourse mental model. In line with these assumptions, also learners benefit from gestures production in learning contexts (Novack & GoldinMeadow, 2015).

Relevant to the present investigation, teachers communicate in classroom settings by using gestures (Alibali et al., 2014) and students can detect conceptual information expressed in those gestures (Kelly & Church, 1998) and can benefit from them (Koumoutsakis et al., 2016). Teacher's gestures have been proved beneficial for learning scientific concepts in instruction done live. For example, a study has revealed improved performance on a posttest after a lesson on the notion of conservation accompanied by representational gestures compared to a lesson that did not contain gestures (Church et al., 2004). Similar results have been obtained by studies concerned with learning math concepts (Valenzo et al., 2003). Also learners' gestures can facilitate learning (Ianì, Cutica & Bucciarelli, 2016). Studies have shown that fourth-grade children who produced gestures during instruction on a math task were more likely to retain and generalize the knowledge they gained, than children who did not gesture (Cook & Goldin-Meadow, 2006). Other studies have revealed that when instructed to gesture while explaining math problems, children added new problem-solving strategies to their repertoire and remembered more from a subsequent lesson from the teacher (Broaders et al., 2007).

Instruction done on video can be as effective as done live (Koumoutsakis et al., 2016) and is seen as a viable alternative to face-to-face teaching for a host of reasons (Allen & Seaman, 2010). Relevant to the aim of our investigation, instruction by video can exploit the power of gestures. For example, studies have revealed that adult learners benefit from videos of teachers gesturing respect to pictures while teaching about gear movement (Carlson et al., 2014) or from videos of teachers pointing at slides while teaching a complex statistical concepts (Rueckert et al., 2017). As for instruction done on video and involving learner's gestures, at our knowledge there are no studies in the literature.

In general, the studies concerning gestures in instruction done live or by video have been concerned with the role of gestures in learning single scientific concepts and it is not obvious that their findings can be extended to learning from a scientific discourse featuring several related concepts. The twofold aim of our investigation was to ascertain whether learning from scientific discourse through videos benefits from gestures, observed or produced. We present three experiments on the role of gestures in learning from science video lessons. In Experiment 1, the participants watched two video lessons: in one video the teacher accompanied the discourse with gestures and in the other video he proffered the discourse while staying still. At recall, the participants recalled few concepts; a possible explanation is the transient high number of concepts. Experiments 2 and 3 used the same scientific lessons but segmented in parts for two reasons. First, the procedure allowed the participants more time to elaborate the information. Second, we could manipulate the variable gesture (no gesture, teacher's gesture, learner's gesture) and gain more in-depth understanding of the effect of gestures on learning from scientific discourse through video lessons.

2 Does Learning from Scientific Video Lessons benefit from Gestures?

A peculiarity of learning from discourse is the transient nature of the information that posits a noteworthy cognitive load on learners. Also instruction by video is transient in that requires learners to keep the disappeared information in mind in order to comprehend the next piece of information (Ayres & Paas, 2007) and this task is demanding for working memory, which has a limited capacity (Cowan, 2001).
Scientific discourse, in particular, has a difficulty that determines an intrinsic load; the higher the number of interacting information elements, the more difficult the material is for the learner and the higher the intrinsic load (Sweller, 1994). Difficulty also depends on learner expertise: with increasing expertise more information elements are combined into schemata, which reduces the intrinsic load of a task.

The participants in our experiment encountered two scientific discourses with a high intrinsic load: they concern topics unknown to the participants (the airplane flying and the sound propagation) and feature a domain-specific terminology. In Experiment 1, the discourses were presented in their normal flow, without interruptions, whereas in Experiments 2 and 3 they were segmented. The rationale was to ascertain whether presenting information in pieces rather than as a continuous stream makes videos more effective for learning (see Spanjers et al., 2012, for this sort of evidence).

2.1 Experiment 1: Science Video Lessons: Teacher's Gestures don't help

The task of the participants in the experiment was watching two videos in which an actor/teacher proffered a scientific discourse; one of the discourse was accompanied by gestures and the other was not. Aim of the experiment was to ascertain whether comprehension and memory was better for the discourse accompanied by gestures (observed-gesture) than for the discourse proffered without gesturing (no-gesture).

Method

Participants. Twenty-eight students from Università di Torino (14 males and 14 females; mean age = 23.9 years, *SD* = 1.6 years) voluntarily took part in the experiment in exchange of course credits and after informed consent.

Material and procedures. The experimental material consisted in two scientific discourses, one concerning the airplane flight (hereafter, Airplane) and the other concerning the nature of sound (hereafter, Sound). The texts were presented in the form of an oral discourse by an actor, each in two conditions: in the observed-gesture condition the actor accompanies the discourse with gestures whereas in the no-gesture condition the actor stays still while proffering the discourse. The actor (the second author, a professional actor with a degree in computer science and qualified to teach in high school) had been instructed to study the two texts before the recording of the videos, and to plan the gestures to produce along with the speech in the observedgesture condition.

Each participant encountered both discourses, one in the observed-gesture condition and the other in the no-gesture condition. Half of the participants dealt first with the observed-gesture condition, and half with the no-gesture condition; in each group, half of the participants encountered first the Airplane discourse, and half of them the Sound discourse. In each condition, the participants attended each discourse twice, and then were invited to recollect as much information as they could. All of the participants were video-recorded.

To code the results, each discourse was divided into 52 semantic units, corresponding to as many main concepts that the learner could recall. Each concept (i.e., semantic unit) recalled by the participants was evaluated by two independent judges according to the following coding schema:

- Correct recollection: a semantic unit recollected literally or as a paraphrase.
- Discourse-based inference: a recollection in which the participant gave explicit information that was originally implicit in the semantic unit.
- Elaborative inference: a semantic unit recollected with the addition of plausible details.
- Erroneous recollection: a recollection with a meaning that was inconsistent with the semantic unit.

Consider, for instance, the following semantic unit in the Airplane discourse: “The air divides as it hits the front of the wing”. According to the coding schema, the statement “(The air) is divided in to two” was a correct recollection; the statement “(The air) coming on the wing separates and for this there are two different speeds” was a discourse-based inference (because it refers to a causal consequent). Now consider the following semantic unit in the discourse: “This difference creates what is known as an aerofoil”, according to the coding schema, the statement “the wing profile is the shape of the wing” was an elaborative inference and the statement “He goes to imagine what a wing profile is” was an erroneous recollection.

Results. The two independent judges coded the participants’ recollections individually. The judges reached a significant level of agreement on their first judgments (agreed on 79.7% of the coding, Cohen’s $K = .56$, $p < .001$). For the final score, the judges discussed each item on which they disagreed, until reaching a full agreement.

Table 1 shows the mean types of recollections in the two conditions of the experiment. Results revealed no differences across the two conditions in the number of erroneous recollections ($t(27)=1.22$, $p = .23$), as well in the number of elaborative inferences ($t(27)=1.66$, $p = .11$) and in the number of discoursed based inferences ($t(27)=1.39$, $p = .18$). Crucially, we did not detect a significant difference across the two conditions in the number of correct recollections ($t(27)=0.21$, $p = .84$).

If the gestures of the speaker were not effective because there was not enough time to elaborate them along with the numerous scientific concepts, then gestures, compared to no gestures, should favor comprehension and learning if the time to process them along with speech is increased. We devised Experiment 2, in which the scientific discourses were presented segmented. Further, in Experiment 2 we added a condition in which listeners were invited to produce their own gestures; since studies in the literature suggest that gestures production can be more effective than gesture observation (e.g., Hornstein & Mulligan, 2004), it is possible that also in case of learning from a scientific discourse gestures production is more effective than gestures observation. Another possibility is that, given the high number of related scientific concepts, learning from a segmented discourse, still transient in nature, does not benefit from learners' gestures.

Table 1. Mean types of recollections (and standard deviations in paretheses) in the no-gesture and observed-gesture conditions in Experiment 1.

Condition	Correct recollections	Discourse-based inferences	Elaborative inferences	Errors
No-gesture	13.80 (4.90)	0.64 (0.95)	0.46 (0.64)	0.68 (1.00)
Observed-gesture	14.00 (5.20)	1.00 (1.20)	0.21 (0.50)	1.00 (1.12)

2.2 Experiment 2: Science Video Lessons: the Learner's Gestures worsen Comprehension and Learning

Each participant in the experiment watched one video in which the actor proffered a discourse. Each video was segmented so that between the introduction of a concept and another one there was a pause. The experiment featured three experimental conditions and each participant was randomly assigned to one of the three conditions: the discourse was proffered by the actor while staying still and the participants attended the video while staying still (no-gesture condition); the actor accompanied the discourse with gestures and the participants stayed still (observed-gesture condition); the actor stayed still and the participants were invited to gesture, in the pause, so to represent the concept just introduced by the actor (produced-gesture condition).

Method

Participants. Thirty students from Università di Torino (3 males and 27 females; mean age = 22.5 years, *SD* = 1.4 years) voluntarily took part in the experiment in exchange of course credits and after informed consent. None of them had taken part in Experiment 1.

Material and procedures. The experimental material consisted in the two scientific discourses of Experiment 1. There were three experimental conditions and each participant was randomly assigned to one of the three: no-gesture, observed-gesture and produced-gesture conditions. In each condition, half of the participants encountered the Airplane discourse and half of them the Sound discourse. Participants in each experimental condition attended the discourse twice, and then were invited to recollect as much information as they could. All of the participants were video-recorded. Participants' recollections were coded as in Experiment 1 by two independent judges.

Results. The two independent judges coded the participants' recollections individually. The judges reached a significant level of agreement on their first judgments (agreed on 89.5% of the coding, Cohen's K = .70, p < .001). For the final score, the judges discussed each item on which they disagreed, until reaching a full agreement.

Table 2 shows the mean types of recollections in the three conditions of the experiment. There was a statistically significant difference between groups in terms of correct recollections as determined by one-way ANOVA ($F(2,27) = 8.12$, $p = .002$, $\eta_p^2 = .38$). A Bonferroni post hoc test revealed that the number of correct recollections was statistically significantly lower in the produced-gesture group compared to the nogesture and the observed-gesture groups (for both comparison, $p = .005$). There was no statistically significant difference between the no-gesture and the produced-gesture groups ($p = 1$). No difference between groups has been detected in terms of discoursebased inferences ($F(2,27) = 0.46$, $p = .64$), elaborative inferences ($F(2,27) = 0.06$, $p = .94$), and erroneous recollections ($F(2,27) = 1.00$, $p = .38$).

Table 2. Mean types of recollections (and standard deviations in parentheses) in the no-gesture, observed-gesture and produced-gesture conditions in Experiment 2.

Condition	Correct recollections	Discourse-based inferences	Elaborative inferences	Errors
No- gesture	21.60 (6.79)	1.00 (1.41)	0.30 (0.48)	1.00 (1.63)
Observed gesture	21.70 (8.31)	1.40 (1.17)	0.30 (0.97)	1.90 (1.25)
Produced gesture	11.30 (4.00)	0.90 (1.10)	0.40 (.97)	1.00 (1.25)

The results of Experiment 2, when compared with those of Experiment 1, suggest that participants in Experiment 1 recalled fewer semantic units because they had difficulty in following the flow of the oral discourse. Participants in Experiment 2 had more time to invest in learning, but evidently not enough to benefit from gestures. Indeed, the results of Experiment 2 have revealed that gestures do not favor learning from a segmented scientific discourse. These results contrast with those of studies in the literature concerning learning from scientific texts; they suggest that environments that allow learner control in terms of going back and forward in the text and in the time invested in learning can exploit the beneficial effect of gestures. For example, studies have revealed that learning from scientific texts benefits from gestures production in learner-paced learning environments (Cutica & Bucciarelli, 2013; Cutica et al., 2014). As for the effectiveness of gestures when observed as compared to when produced, the results of Experiment 2 reveal that gesture production worsen learning. Experiment 3 was a within-subjects study whose aim was to investigate more in-depth the effects of gestures observation and gestures production in learning from scientific video lessons.

2.3 Experiment 3: Science Video Lessons: A more in depth investigation into the effect of Teacher's and Learner's gestures

The experiment was a replication of Experiment 2, but the design was within and there were only two conditions: observed-gesture and produced-gesture.

Method

Participants. Twenty students from Università di Torino (4 males and 16 females; mean age = 22.6 years, *SD* = 4.8 years) voluntarily took part in the experiment in exchange of course credits and after informed consent. None of them had taken part in Experiments 1 and 2.

Material and procedures. The experimental material was the same as for Experiments 1 and 2. There were two experimental conditions: observed-gesture and producedgesture. Each participant dealt with the two conditions: half of them encountered the Airplane discourse in the observed-gesture condition and the Sound discourse in the produced-gesture condition, and half of them encountered the Sound discourse in the observed-gesture condition and the Airplane discourse in the produced-gesture condition. The experimental procedures and the coding of results was also the same as for the previous experiments.

Results. Two independent judges coded the participants' recollections individually. The judges reached a significant level of agreement on their first judgments (agreed on 94.6% of the coding, Cohen's $K = .75$, $p < .001$). For the final score, the judges discussed each item on which they disagreed, until reaching a full agreement.

Table 3 shows the mean types of recollections in the two conditions of the experiment. We detected a significant difference across the two conditions in terms of correct recollections ($t(19) = 3.01$, $p = .007$, d'= 0.676). No difference was detected for elaborative inferences ($t(27)=1.00$, $p = .33$), discoursed based inferences ($t(27)=0.28$, $p = .82$) and erroneous recollections ($t(27)=0.34$, $p = .74$).

Table 3. Mean types of recollections (and standard deviations in parentheses) in the observedgesture and produced-gesture conditions in Experiment 3.

Condition	Correct recollections	Discourse-based inferences	Elaborative inferences	Errors
Observedgesture	21.20 (6.18)	0.70 (1.17)	0.05 (0.22)	0.95 (1.28)
Producedgesture	16.20 (6.43)	0.60 (1.27)	0.35 (1.57)	0.80 (1.32)

3 Discussion and Conclusions

The aim of the present investigation was to explore the effect of gestures accompanying a scientific discourse delivered through video. In three experiments, we ascertained whether learning from scientific discourse through videos benefits from gestures, observed or produced. Globally considered, the results of Experiments 1 and 2 have revealed that comprehension and learning from a scientific discourse do not improve when the teacher gestures compared to when the teacher does not gesture, and the results of Experiment 2 and 3 have revealed that the learner's gestures, as compared to the teacher's gestures, can worsen comprehension and learning.

A tentative explanation for this global pattern of result is that since science lessons involve knowledge of a high specialized vocabulary, when learners lack this knowledge, they might struggle to construct the discourse mental model also when the discourse is accompanied by gestures. Indeed, within a mental model perspective, a text can be represented at three levels: the surface representation, the textbase representation, and the mental model representation. The surface representation consists of the verbatim words and clauses extracted from the discourse. At the textbase level, the meanings of words and clauses are processed and subsequently stored in the listener's memory. A mental model representation is a coherent and non-linguistic mental representation of the 'state-of-affairs' described in a discourse, whose precondition is the processing of the meanings of words.

But why gestures produced as compared to gestures observed do worsen comprehension and learning? A possible explanation is that the high specialized language involved in our scientific discourse posited a high cognitive load on participants and the worry to plan and produce gestures was an extra-demand for participants in the produced-gesture condition as compared to observed-gesture condition.

A significant difference between Experiment 1 and Experiments 2 and 3 is that the latter used segmented video lessons. Relevant for the present investigation, segmented video lessons, compared to video lessons presented in their natural flow, favored comprehension and learning. This result is in line with those of studies revealing that segmenting information, that is presenting them in pieces rather than as a continuous stream, makes videos more effective for learning in that have a beneficial effect on cognitive load and learning for novices. Pauses inserted between the segments may give learners extra time to perform necessary cognitive processes (Spanjers et al., 2012). By segmenting the discourse and providing pauses between a segment and another, participants could have the time to apply more elaborate discourse-processing strategies.

In conclusion, the implications of the results of the three experiments for technology enhanced learning are that video lessons in which the teacher gestures seem to have no advantage on video lessons in which the teacher does not gestures, but segmented video lessons should be preferred to video lessons delivered in their plain flow.

References

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Johnson Martinez, C., … Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction, 32*(1), 65–100.

Allen, I. E., & Seaman, J. (2010). *Learning on demand: Online education in the United States, 2009*. Newburyport, MA: The Sloan Consortium.

Ayres, P., & Paas, F. (2007). Can the cognitive load approach make instructional animations more effective? *Applied Cognitive Psychology, 21*, 811-820.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*, 539-550.

Bucciarelli, M. (2007). How the construction of mental models improves learning. *Mind & Society*, *1*(6), 67-89.

Carlson, C., Jacobs, A. A., Perry & Brecknridge Church, R. (2014). The effect of gestured instruction on the learning of physical causality problems. *Gesture*, *14*(1), 26-45.

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7*(4), 303-319.

Cook, S.W., & Goldin-Meadow, S. (2006). The role of gestures in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7*, 211–232.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science*, *24*, 87-185.

Cutica, I., & Bucciarelli, M. (2008). The deep versus the shallow: Effects of co-speech gestures in learning from discourse. *Cognitive Science*, *32*, 921-935.

Cutica, I. & Bucciarelli, M. (2011). "The more you gesture, the less I gesture": Cospeech gestures as a measure of mental model quality. *Journal of Nonverbal Behavior*, 35, 173-187.

Cutica, I. & Bucciarelli, M. (2013). Cognitive change in learning from text: Gesturing enhances the construction of the text mental model. *Journal of Cognitive Psychology*, *25,2*, 201-209.

Cutica, I. & Bucciarelli, M. (2015). Non-determinism in the uptake of gestural information. *Journal of Nonverbal Behavior*, *39*, 289-315

Cutica, I., Ianì, F., & Bucciarelli, M. (2014). Learning from text benefits from enactment. *Memory & Cognition*, *42*, 1026-1037.

Hornstein, S. L., & Mulligan, N. W. (2004). Memory for actions: Enactment and source memory. *Psychonomic Bulletin & Review*, *11*(2), 367-372.

Ianì, F., Cutica, I., & Bucciarelli, M. (2016). Timing of gestures: Gestures anticipating or simultaneous with speech as indexes of text comprehension in children and adults. *Cognitive Science*, *41*(6), 1549-1566.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, and consciousness*. Cambridge University Press, Cambridge, UK.

Kelly, S. & Church, R.B. (1998). A comparison between children's and adults' ability to detect conceptual information conveyed through representational gestures. *Child Development*, *69*, 85-93.

Koumoutsakis, T., Breckindridge Church, R., Alibali, M. W., Singer, M., & AymanNolley, S. (2016). Gesture in instruction: Evidence from live and video lessons. *Journal of Nonverbal Behavior*, *40*, 301-315.

Novack, M., & Goldin-Meadow, S. (2015). Learning from gesture: How our hands change our minds. *Educational Psychology Review*, *27*(3), 405-412.

Rueckert, L., Breckinridge Church, R., Avila, A., & Trejo, T. (2017). Gesture enhances learning of a complex statistical concept. *Cognitive Research: Principles and Implications*, *2*:2.

Spanjers, I. A. E., Van Gog, T., Wouters, P., & Van Merriënboer, J. J. G. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, *59*, 274-280.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295-312.

Valenzano, L, Alibali, M. W., & Klatzky, R. (2013). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187-204. van Dijk, I.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Zwaan, R. A., & Radvansky G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123(2),* 162-185.