






# How Cross-Representational Signaling Affects Learning from Text and Picture: An Eye-Tracking Study

Juliette C. Désiron<sup>1</sup> , Mireille Bétrancourt<sup>1</sup> ,  
and Erica de Vries<sup>2</sup> 

<sup>1</sup> University of Geneva, Geneva, Switzerland  
{Juliette.Desiron,Mireille.Betrancourt}@unige.ch

<sup>2</sup> Univ. Grenoble Alpes, Grenoble, France  
erica.devries@univ-grenoble-alpes.fr

**Abstract.** Multimedia learning research pointed out that adding a picture to a text is not systematically beneficial to learners. One of the most influential factors is the necessity for learners to identify mutually referring information in the written and pictorial representations. This study investigates how Cross-Representational Signaling (CRS) facilitates learning from multimedia document. In this study, CRS is implemented by mutually referring visual and verbal cues which highlight semantic links between text and picture. Two versions of the same multimedia document explaining the risks of being caught in a rapid, with or without CRS, are compared. The study that is still ongoing will provide data on online processing (eye-tracking data) and learning outcomes. The results will provide insights on the use of CRS to improve the design of instructional diagrams.

**Keywords:** Multimedia learning · Eye-tracking · Signaling  
Cross-representational signaling

## 1 Theoretical Framework

Diagrams and pictorial representations are often used to support comprehension of instructional documents. Multimedia learning research showed that learning with multiple representations (particularly written text and pictures) can be beneficial to comprehension provided that learners can identify links between representations through cross-references [3, 10]. The most widely accepted models of multimedia learning (CTML from Mayer [6]; ITPC from Schnotz and Bannert [10]) claim that information from verbal and pictorial representations are first processed by media specific (verbal or pictorial) channels before being integrated in a coherent model of the situation relying on both those representations and previous knowledge. The latest version of the ITPC model from Schnotz [9] includes a coherence principle which predicts that “students learn better from words and pictures than from words alone if the words and pictures are semantically related to each other” ([9] p. 23), especially for students with poor reading skills or little prior knowledge.

---

The original version of this chapter was revised: The publishing mode has been changed to open access. The correction to this chapter is available at [https://doi.org/10.1007/978-3-319-91376-6\\_82](https://doi.org/10.1007/978-3-319-91376-6_82)

Effectively guiding learners' integration processes may be channelled through the insertion of visual or verbal cues in either one or both verbal and pictorial representations [11]. A meta-analysis by Richter et al. [8] found an overall significant beneficial effect of signaling text-picture relations on comprehension that was more profitable to low to medium prior-knowledge than to high prior-knowledge learners. These results confirm the ITPC [9] claim that supporting text-picture semantic links facilitates the construction of a coherent mental representation. A possible moderating effect of reading abilities was however not investigated.

Kalyuga et al. ([4] Exp. 2) used interactive colour coding in both representations to facilitate search of corresponding verbal and pictorial elements. The cueing group performed significantly better than the no-cueing group. Using eye-tracking to compare the use of verbal cues (labelling) in the pictorial representation Mason et al. [5] found that more integrative processing, measured through eye-fixations, occurred with labelled pictures. This research shows that eye-tracking data can give interesting insights in on-line processes of text-picture integration.

In the present study, we implemented Cross-Representational Signaling (CRS) through colour coding cues and picture labelling to highlight semantic links between written texts and visual pictures. Two versions (with or without CRS) of the same multimedia document (a 5-page text and picture instruction explaining the risks of being caught in a rapid) were designed. After completing reading skills tests, participants learned with one of the two versions of the multimedia document and answered comprehension (text-based and inference) questions.

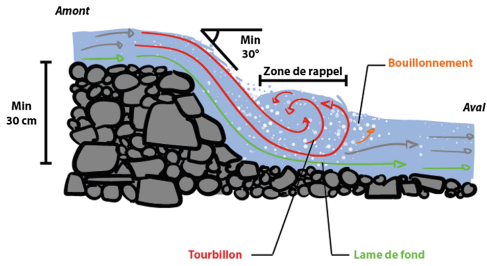
We assume that CRS facilitates the construction of a coherent mental model, which should lead to better comprehension scores, especially for students with lower reading skills. Eye-tracking data will provide insights on the way CRS affects the processing of instructional diagrams. In particular, following Mason et al. [5], we expect that signaling in the text will prompt exploration of the pictures and increase the total time spent on the pictures.

## 2 Method

The experimental material was a 5-page expository document including text and static representational pictures on how to escape the Maytag effect when being caught in a rapid. The material was carefully selected and designed to ensure that both media were necessary for no prior knowledge learners to comprehend the document. The pictures were designed to be representational in the sense of Carney and Levin [1]. CRS encompassed the following mutually referencing verbal and visual cues: colours, symbols and labels (see Figs. 1 and 2). The material was presented on a 23" screen, and participants' eye-movements are recorded with a Tobii TX300.

Participants' prior knowledge was evaluated online, before the experiment, with a self-assessed multiple-choice knowledge questionnaire. Because the ITPC advocates that semantic links are helpful only to low prior knowledge readers, only participants with low or no prior knowledge on the topic were recruited. During the experiment, participants completed two reading skills assessments (a vocabulary test from Deltour [2] and an inference generation test adapted from Meteyard et al. [7]). Then they studied the multimedia document in one of two experimental conditions (with vs. without CRS). After reading, participants completed a 7 items Likert scale

[...]  
Formée en **amont** (avant), la **lame d'eau de profondeur**, échappe toutefois au **tourbillon** pour ressurgir en **aval** (après) d'un **bouillonnement d'eau**. Dans celui-ci la densité de l'eau est divisée par 2, ce qui limite la flottaison.



**Fig. 1.** Sample of the material with CRS

[...]  
Formée en **amont** (avant), la lame d'eau de profondeur, échappe toutefois au tourbillon pour ressurgir en **aval** (après) d'un bouillonnement d'eau. Dans celui-ci la densité de l'eau est divisée par 2, ce qui limite la flottaison.



**Fig. 2.** Sample of the material without CRS

questionnaire on motivation, perceived difficulty and perceived effort. They ended with the comprehension test, with 13 open-ended questions at three levels: text-base comprehension, local bridging and global-bridging inference. A drawing task was also included, in which participants had to draw and name the different currents involved in the formation of a whitewater.

The experiment was still running when we wrote this paper. A random sample of 40 to 50 undergraduate university students in education sciences or psychology will be recruited overall.

### 3 Data Analyses and Expected Results

Following previous research in multimedia learning using eye-tracking as an online measure of comprehension [3, 5], we will analyze the collected data with first-pass and second-pass fixations. Specifically, we will consider fixations as gazes and focus on look from text to picture, both in general and with targeted AOI.

First, following the ITPC model [9] we expect that multimedia comprehension will be higher in the CRS than in the control condition, especially for students with low reading skills. Regarding on-line processing, we expect that participants reading the multimedia document with CRS will look at the picture during first-pass and second-pass reading more often than participants without CRS. Indeed, research by Mason et al. [5] pointed out that a picture with verbal cues elicited more integration with the text than a picture without verbal cues. Further exploratory analyses of eye-tracking data will provide insights on how text-picture integration processes differ with and without CRL. Participants reading skills will be inserted in the analyses as a potential moderator.

This study will contribute to test an implementation of the *coherence* condition, theoretically developed in the ITPC model, when a document is designed with

Cross-Representational Signaling. The findings will provide guidelines regarding the design of commented diagrams used for instructional or public awareness purposes.

**Acknowledgments.** This study was supported by the Swiss National Science Foundation with a Doc.CH grant attributed to the first author [POGEP1\_165256].

## References

1. Carney, R.N., Levin, J.R.: Pictorial illustrations still improve students' learning from text. *Educ. Psychol. Rev.* **14**(1), 5–26 (2002). <https://doi.org/10.1023/A:1013176309260>
2. Deltour, J.J.: Echelle de vocabulaire de Mill Hill de JC Raven: Adaptation française. In *Manuel des Raven*. Braine le Château. Oxford Psychologists Press, Belgique (1993)
3. Hegarty, M., Just, M.A.: Constructing mental models of machines from text and diagrams. *J. Mem. Lang.* **32**, 717–742 (1993). <https://doi.org/10.1006/jmla.1993.1036>
4. Kalyuga, S., Chandler, P., Sweller, J.: Managing split-attention and redundancy in multimedia instruction. *Appl. Cogn. Psychol.* **25**(Suppl. 1), 351–371 (1999). <https://doi.org/10.1002/acp.1773>
5. Mason, L., Pluchino, P., Tornatora, M.C.: Effects of picture labeling on science text processing and learning: evidence from eye movements. *Read. Res. Q.* **48**, 199–214 (2013). <https://doi.org/10.1002/rq.41>
6. Mayer, R.E.: Cognitive theory of multimedia learning. In: Mayer, R.E. (ed.) *The Cambridge Handbook of Multimedia Learning*, pp. 31–46. Cambridge University Press, Cambridge (2005). <https://doi.org/10.1017/CBO9780511816819.004>
7. Meteyard, L., Bruce, C., Edmundson, A., Oakhill, J.: Profiling text comprehension impairments in aphasia. *Aphasiology* **29**(1), 1–28 (2015). <https://doi.org/10.1080/02687038.2014.955388>
8. Richter, J., Scheiter, K., Eitel, A.: Signaling text-picture relations in multimedia learning: a comprehensive meta-analysis. *Educ. Res. Rev.* **17**, 19–36 (2016). <https://doi.org/10.1016/j.edurev.2015.12.003>
9. Schnotz, W.: Integrated model of text and picture comprehension. In: Mayer, R.E. (ed.) *The Cambridge Handbook of Multimedia Learning*, 2nd edn., pp. 72–103. Cambridge University Press, New York (2014). <https://doi.org/10.1017/cbo9780511816819.005>
10. Schnotz, W., Bannert, M.: Construction and interference in learning from multiple representation. *Learn. Instr.* **13**, 141–156 (2003). [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
11. van Gog, T.: The signaling (or cueing) principle in multimedia learning. In: Mayer, R.E. (ed.) *The Cambridge Handbook of Multimedia Learning*, 2nd edn., pp. 263–278. Cambridge University Press (2014). <https://doi.org/10.1017/cbo9781139547369.014>

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

