Commentary on "COGNITIVE EFFECTS OF ARGUMENT VISUALIZATION TOOLS" by Michael Hoffmann

FABIO PAGLIERI

Goal-Oriented Agents Laboratory (GOAL)
Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche (ISTC-CNR)
Via S. Martino della Battaglia 44, 00185 Roma
Italy
fabio.paglieri@istc.cnr.it

Ergonomists have often meditated about a little game known as "the game of 15": two players in turn say a number between 1 and 9, and each particular number may not be repeated; the game is won by the player who has first said three numbers whose sum is 15; if all numbers are exhausted and no combination of three numbers said by either player adds up to 15, the game is tied. These rules are simple enough, but the resulting game is not. To play it skilfully, each player needs to remember what numbers have been already said and mentally calculate all possible winning combinations, both to achieve one of them and to prevent the opponent from doing the same. Even good reasoners find it extremely difficult to master this game without the aid of pencil and paper, and such aids improve performance but do not make the game easier to play. As a result, the game of 15 is rarely played, if ever.

Interestingly, there is another game which is isomorphic to the game of 15 and is extremely popular all around the world: Tic-Tac-Toe, also known as Noughts and Crosses. Figure 1 illustrates the isomorphism between the two games: there are exactly eight triplets of numbers between 1 and 9 whose sum is 15, and they correspond to the eight winning combinations on the 3x3 matrix of Tic-Tac-Toe (three rows, three columns, and two diagonals). To put a mark on a square of the Tic-Tac-Toe matrix is equivalent to say a number in the game of 15, thus preventing the opponent to occupy the same square / say the same number and at the same time gaining access to some potential strategies for victory. However, contrary to the game of 15, Tic-Tac-Toe is extremely easy to play effectively, mostly because it represents the same reasoning task in a format which allows to offload all the irrelevant cognitive workload and concentrate on the strategic features of the game (for extensive discussion of this case, see Norman, 1993).

8	3	4
1	5	9
6	7	2

Fig. 1. Isomorphism between the game of 15 and Tic-Tac-Toe

Cognitive offloading is also the key concern of Hoffmann's article: the author is interested to assess to what extent two different Computer-Supported Argument Visualization

Zenker, F. (ed.). *Argumentation: Cognition & Community. Proceedings of the 9th International Conference of the Ontario Society for the Study of Argumentation (OSSA), May 18-21, 2011*. Windsor, ON (CD ROM), pp. 1-4.

(CSAV) tools, Rationale™ (http://rationale.austhink.com/) and AGORA (http://agora.gatech.edu/), effectively reduce the cognitive load of the users, thus facilitating the achievement of their educational aim—to wit, improving argument evaluation and critical thinking. Hoffmann argues that AGORA is superior to Rationale™ in terms of cognitive offloading: the cognitive load that a learner has to handle to construct a good argument is more effectively delegated to the software in AGORA than what happens in Rationale™. Since Hoffmann is the leader of the AGORA project, his endorsement of this software is hardly surprising: however, the paper presents objective arguments to compare the two CSAV tools, and I see no reason to suspect any bias or hidden agenda in Hoffmann's line of reasoning.

Overall, I am inclined to agree with Hoffmann's conclusion: it would seem indeed that AGORA permits greater cognitive offloading than Rationale™, thus helping users to focus more on the relevant task supported by these tools, i.e. argument evaluation. However, there are two important considerations that limit the relevance of such conclusion: first, the greater cognitive load required to use Rationale™ could well be the result of its greater expressivity, in comparison to AGORA; second, effective cognitive offloading in and by itself is not sufficient to pass judgment on rival CSAV tools, for reasons different from those considered by Hoffmann.

Let us start with the first point: as Hoffmann admits, Rationale™ is designed to support the construction and analysis of a broad family of arguments, whereas AGORA focuses only on deductively valid arguments. Hence Rationale™ needs to define a rather general and loose system of rules, such that any kind of intuitively good argument can be represented within it, regardless of the pertinent standard of validity for that particular case (deductive, inductive, presumptive, etc.). In contrast, AGORA is dedicated only to represent deductively valid arguments, and it limits to seven the admissible argument schemes: *modus ponens*, *modus tollens*, disjunctive syllogism, not-both syllogism, equivalence, conditional syllogism, and constructive dilemma. Given its narrower scope, it is not surprising that AGORA can make use of more restrictive and unambiguous rules than Rationale™. In particular, much of Hoffmann's criticism of Rationale™ hinges upon the fact that its rules are cumbersome to handle for users and also present potential loopholes in terms of argument quality: as Hoffmann notes, in Rationale™ «arguments of different (or dubious) quality can be constructed which nevertheless fulfill the Rabbit Rule and the Holding Hands Rule» (p. 10). This is certainly true, but it seems to be the price one has to pay in order to have a more flexible CSAV tool, one that is not restricted to model deductively valid arguments. At least, such trade-off between cognitive offloading and expressivity should be considered, when comparing different CSAV software.

More generally, drawing a comparison between two CSAV tools that are intended to analyze different sets of arguments (every intuitively good argument for Rationale™, only those that are deductively valid in the case of AGORA) risks confounding two different issues: the effectiveness of each tool, given its intended target, and the instructional legitimacy of such target. Hoffmann here aims to discuss the former, but the latter crops into the discussion anyway. Should we accept the more constraining rules of AGORA, in order to benefit from its effective cognitive offloading? Well, it ultimately depends on whether or not we are happy to work with a CSAV tool that only processes deductively valid arguments. There might well be contexts where this is an excellent option, for instance to familiarize students with real life examples of logical reasoning, as a

complement to an introductory course in logic. But in other contexts this limitation may be highly problematic, since even *good* everyday reasoning is rarely deductively valid (Paglieri & Woods 2011), thereby exceeding the expressive capacity of AGORA: so, for instance, this software would be rarely of practical use to help students reconstructing arguments in political debates or newspaper articles.

This observation leads us to the second limitation of Hoffmann's conclusion: the fact that AGORA allows an effective cognitive offloading is not sufficient to consider it superior to Rationale™ as a CSAV tool. Hoffmann agrees with this claim, but for reasons different from those I will develop here. According to Hoffmann, "Cognitive Load Theory cannot be used to show the superiority of one CSAV tool over another since the actual cognitive load always depends on the expertise of the user" (p. 11). That is correct, but it is not such a formidable objection: after all, it can be circumvented by making one's conclusion equally context-dependent. This is indeed Hoffmann's strategy, when he speculates that

> … for learning the first steps of critical thinking—that is, to get a feeling of what an argument *is*, its structure, and its quality—it might be better to start with a system that provides more guidance and that allows the bootstrapping of those skills that can then be used for many other important things. (p. 11)

In other words, the more limiting environment provided by AGORA could well be more adequate for absolute beginners, who would have serious troubles in handling the more open-ended and cognitively demanding rules of Rationale™.

Fair enough. But the trouble here is that we do not know whether either AGORA or Rationale™ are in fact effective in pursuing their educational aims, regardless of the cognitive load they impose on users. As Hoffmann clarifies, the main function of these CSAV tools is to enable students to assess the quality of arguments; subordinately, they also aim to enable students to identify hidden assumptions in arguments (p. 6). The obvious question is whether either software is minimally successful in fostering these learning objectives, and according to what standard or metrics of evaluation: however, Hoffmann's article presents no evidence of that. Without such evidence, any analysis of the cognitive load of these tools can tell us only half of the story: it could well be the case that AGORA is cognitively more cost-effective than Rationale™, as Hoffmann argues, but this is relevant only insofar as AGORA is also at least equally successful as Rationale™ in yielding instructional benefits to its users. This is indeed the cornerstone of any cost-benefit analysis of alternative software solutions: minimizing costs is important only if it does not jeopardize too much expected benefits.

Of course, assessing the efficacy of CSAV tools in achieving instructional aims requires empirical studies, most notably field studies on populations of users across different educational contexts. This goes well beyond the stated purposes of Hoffmann's paper, who is concerned here with purely theoretical considerations on cognitive load (p. 3). As a consequence, his analysis defines an interesting starting point in a very important research direction, but further studies will be needed to really assess the comparative merits and shortcomings of different CSAV tools.

Incidentally, it is worth emphasizing that such comparison is now particularly urgent, so that Hoffmann is to be commended in drawing our attention to it. CSAV tools are proliferating, and it is paramount to get a better a understanding of whether they work or not, and to what extent and for what purposes—even more so, since the evidence of

their efficacy reviewed by Hoffmann is patched and conflicting (Carr 2003, Bell 2004, Twardy 2004, van Gelder et al. 2004). In order to perform a thorough evaluation, we will also have to pay more attention to specific details and concrete contexts: for instance, it is somehow disappointing that no description of the intended users of either AGORA or Rationale™ is provided in the paper. This is not Hoffmann's fault, but rather reflects the implicit universalistic ambition of such tools: even if they tend to be used mostly within University courses, in principle they are conceived as supporting argument assessment for all kinds of users—from children in primary school to adults. I believe this lack of specificity to be misguided: CSAV tools specifically tailored for the educational needs and competences of different populations of users would, to my mind, prove more effective in bootstrapping argumentation skills. Testing this hypothesis will of course require empirical studies: in fact, these are now much more relevant to establish the validity of CSAV tools than any armchair speculation on their features—even when such speculation is carried out with great diligence and insight, as in the case of Hoffmann's paper.

REFERENCES

Bell, P. (2004). Promoting students' argument construction and collaborative debate in the science classroom. In: Linn, M.C., Davis, E.A., and Bell, P. (eds). *Internet environments for science education* (pp. 115-143). Mahwah, NJ: Erlbaum.

Carr, C.S. (2003). Using computer supported argument visualization to teach legal argumentation. In: Kirschner, P.A., Buckingham Shum, S.J., and Carr, C.S. (eds). *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 75-96). London: Springer.

Gelder, T. van, Bissett, M., and Cumming, G. (2004). Cultivating expertise in informal reasoning. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale* 58(2), 142-152.

Hoffmann, M. (this volume). Cognitive effects of argument visualization tools. In F. Zenker (ed.), *Argumentation: Cognition & Community. Proceedings of the 9th International Conference of the Ontario Society for the Study of Argumentation (OSSA)* (pp. 1-12). Windsor, ON (CD ROM).

Norman, D. (1993). *Things that make us smart*. New York: Doubleday/Currency.

Paglieri, F., and Woods, J. (2011). Enthymematic parsimony. *Synthese* 178(3), 461-501.

Twardy, C. R. (2004). Argument maps improve critical thinking. *Teaching Philosophy* 27(2), 95-116.