COGNITIVE IMPACT METRICS: APPLYING MACROCOGNITION DURING THE DESIGN OF COMPLEX COGNITIVE SYSTEMS

Brian M. Moon
Stephanie Wei
Donald A. Cox
Klein Associates Inc.
Fairborn, Ohio

Macrocognition is an emerging theoretical and methodological framework for describing cognitive work as it naturally occurs (Klein, Ross et al., 2003). It can form the basis for the design of complex cognitive systems that augment, rather than degrade, proficient performance. This paper presents a method for using macrocognition during design to anticipate how a complex cognitive system will impact cognition. We have developed a suite of metrics we call "Cognitive Impact Metrics" (CIM). Because they highlight the necessary features of and potential barriers to proficient cognitive performance, these metrics and their associated measures provide us with a framework in which we can generate predictions about where and how our system will enhance or hinder our performance. Application of CIM may be particularly useful in the design of systems where many potential applications must be culled down to a more manageable set of candidates.

SUMMARY

Introduction

Macrocognition is an emerging theoretical and methodological framework for describing cognitive work as it naturally occurs (Klein, Ross et al., 2003). It focuses practitioners who describe cognition in natural settings on key functions that are emergent and fluid. Thus, these descriptions can form the basis for the design of complex cognitive systems in which information technologies are intended to augment proficient performance.

Our current list of the major macrocognitive functions appears in the center of Figure 1. The circle around these primary functions shows a range of supporting macrocognitive processes, which decision makers carry out as a means for achieving the primary functions listed. The functions, however, emerge repeatedly as ends in themselves across a variety of work in various domains. One of the primary reasons we propose the framework is to encourage research on these otherwise ignored, but highly important phenomena (Klein, Ross et al., 2003).

Thesis

This paper will suggest a method for using descriptions of macrocognitive functions and processes as the basis for anticipating how a complex cognitive system can succeed or fail in providing cognitive support. We call this suite of descriptions *Cognitive Impact Metrics (CIM)*. The individual metrics define what we want to examine. The measure associated with a metric is how we make judgments about the value of a particular system element, be it a requirement, design concept, feature, or application. Because they highlight the necessary features of and potential barriers to proficient cognitive performance, these metrics and their associated measures provide us with a framework through which we can anticipate where and how system elements will enable or hinder cognitive activity. Within a particular situation or context, we can use the metrics and measures to generate a "gold standard" for support and performance. Deviations from this standard, then, can serve as a means of rating system elements, and therefore we can use CIM to compare amongst alternative elements for potential inclusion in the overall system design.

Macrocognition

Ross et al. (2002) provide high-level descriptions, but an in-depth understanding of the macrocognitive processes and functions is required to generate metrics. There may be a steep learning curve to understanding the meaning and importance of each measurement. We have drawn on the in-depth descriptions captured in the references listed below.

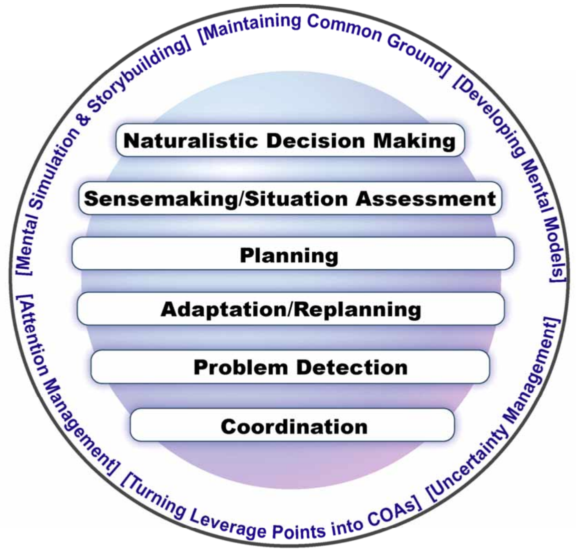

Figure 1. Macrocognition.

As Klein Associates' work has expanded, a set of cognitive functions and processes has been found to underlie effective human cognitive performance. Termed macrocognition, we are developing within this framework models to describe and solutions to support effective cognitive performance.

Generating CIM

The CIM were generated starting with a review of the existing models of the macrocognition framework. The description of each model and its role in effective performance was examined to locate situations or behaviors that could be observed, evaluated, and compared. The goal for the resulting list of metrics and measures is to be a useful and usable guide to the practitioner and to capture essential features of the relevant model. We have searched for measures that can be used to evaluate, and generate a standard against which we can compare the proposed system.

It is worth noting that we currently do not have CIM for every macrocognitive function and process, primarily because the descriptions and models for some are immature or nonexistent. We have developed, or are developing, CIM for:

- Sensemaking (Klein, Phillips, Battaglia, Wiggins, & Ross, 2002),
- Planning (Thunholm, 2003),
- Replanning (Klein, Wiggins, & Lewis, 2003; Klein, Wiggins, & Schmitt, 1999),
- Problem Detection (Klein, Pliske, Crandall, & Woods, in press),
- Maintaining Common Ground (Klein, Armstrong, Woods, Gokulachandra, & Klein, 2000), and
- Uncertainty Management, Analytic Decision Making (Klein, 2003).

Tables 1-2 provide descriptions for two of the existing sets of CIM and their measures. In our current use the measures are independent in that there is no rule or formula to compute the score for a metric depending on the rating of the contained measures.

Evaluation Using CIM

Someone evaluating an application will want to know the full range of its functions and consider how each lines up against his or her measures. Where applications provide support for measures—through features, combinations of features, and envisioned uses—their ratings can be commensurate with their level of support. Where they disable, obstruct, interrupt, or otherwise hinder the measures, their ratings can reflect these disadvantages. In this way, CIM form the basis for a Consumer Reports©-type assessment — alternative system elements are compared using a set of standard criteria. Analysis using the CIM should focus on the degree to which the proposed intervention stands up to the standard. "Degree" can refer to both a quantitative and a qualitative measure; both can be used in an analysis. Quantitatively, the analyst may want to create a scoring system for the measures, awarding and subtracting points depending on whether or not a given measure is supporting or detracting the user of the application. Qualitatively, the analyst may make judgments against the measure given the "weight" of the evidence. A three-level scale is probably most useful here to indicate where there is a clear support or hinder call or a middle ground where the evidence may not be as incontestable. A contestable rating may suggest additional data collection is necessary.

Care must be taken to document the process and rationale by which practitioners select their metrics and make their ratings. A comprehensive, yet understandable, presentation of the analysis is critical to its gaining acceptance and application by system designers.

Application to System Development

Human Factors practitioners are called upon to evaluate systems at many points in the development process. The nature and description of the system that they must use to perform the evaluation may range from a proposal a few pages long to a multi-volume requirements document to a collection of COTS applications and a vague notion of a new way of working. The CIM are an attempt to bring the macrocognitive perspective into the system development process at all stages, under any conditions. Malek, Alvidrez, Moon, and Wei (2004) suggest that repeatable, predictable engineering of effective systems will only occur when the understanding of cognitive functions of job performance is integral to the design and evaluation of the system.

As with other evaluation frameworks, the CIM suite can help inform the design before it is formally evaluated. In envisioning how well an intervention might support macrocognition, designers consider different arrangements and applications of technologies, or perhaps new tactics, techniques, and procedures that, when properly implemented, can provide even greater support for the processes and functions.

For example, it is generally accepted in the development community that system quality is enhanced by having reviews and inspections (Freedman & Weinberg, 1990) at all stages of the development process. This is especially true before the system has reached the point in maturity when empirical data can be gathered. It is highly desirable to detect problems as early as possible in the development process, as they become more costly to fix the later they are found.

The use of different perspectives has been shown to increase the quality of inspections and expert reviews (Basili et al., 1996). The CIM are used to add another perspective to the review process. A HF practitioner familiar with the macrocognitive framework and its constituents may be able to cover the whole CIM suite in a single review. For additional coverage, or for the less experienced, focusing on only a few of the macrocognitive functions and processes per reviewer or review pass.

Taking a risk-management view of the development process (DeMarco & Lister, 2003), we can see that the CIM provide a means for making visible, and hence managing, the risk of producing a system that interferes with the ability of the people in and around the system to use their macrocognitive functions and processes to full positive effect. Expert evaluations early in the development process will not guarantee the performance of the final system, but they will allow development managers and HF practitioners to guide the system design and development toward one that allows people to perform at their peak capacity**.**

Our goal with Cognitive Impact Metrics is to provide HF practitioners, system designers and program managers with a reliable, usable, and manageable set of standards for anticipating cognitive impact of technologies. One of the strengths we see for CIM is that they can be applied in a wide variety of situations, whether developing a complex real-time command-and-control system from the ground up or a system made of a suite of existing commercial applications.

Future Development

The models that make up the macrocognition framework are continuing to evolve as they are applied in new domains and are used more widely. The CIM will evolve to track the models as they change. An early version of the CIM suite was used to evaluate a suite of existing applications as a part of a work process redesign effort. As the CIM suite are applied in a wider range of system development processes and contexts, we expect to further refine how the CIM are used and how they are presented.

ACKNOWLEDGEMENT

The research reported here was supported by Contract MDA972-03-9-0001 prepared for DARPA, Subcontract Hicks and Associates 4400062705.

REFERENCES

Basili, V. R., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Sorumgard, S., Zelkowitz, M. V. (1996). The empirical investigation of perspective-based reading. *Empirical software engineering,* 1, 133-164. Boston: MA. Kluwer Academic Publishers.

DeMarco, T. & Lister, T. (2003) *Waltzing with bears: Managing risk on software projects.* New York: Dorset House.

Freedman, D. P., & Weinberg, G. M. (1990) *Handbook of Walkthroughs, Inspections, and technical Reviews.* New York: Dorset House.

Klein, G. (2003). *Intuition at work*. New York: Doubleday.

Klein, G., Armstrong, A., Woods, D. D., Gokulachandra, M., & Klein, H. A. (2000). *Cognitive wavelength: The role of common ground in distributed replanning* (Final Technical No. Report No. AFRL-HE-WP-TR-2001-0029). Wright-Patterson AFB, OH: United States Air Force Research Laboratory.

Klein, G., Phillips, J. K., Battaglia, D. A., Wiggins, S. L., & Ross, K. G. (2002). *Focus: A model of sensemaking* (Interim Report - Year 1 prepared Alexandria, VA). Fairborn, OH: Klein Associates.
Klein, G., Pliske, R. M., Crandall, B., & Woods, D. (in press). Problem detection. *Cognition, Technology, and Work.*
Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems, 18*(3), 81-85.
Klein, G., Wiggins, S. L., & Lewis, W. R. (2003). Replanning in the Army brigade command post. In *Proceedings of the 2003 CTA Symposium [CD-ROM].*
Klein, G., Wiggins, S. L., & Schmitt, J. (1999). *Cognitive aspects of replanning at Army division-level command posts* (Final Report No. Contract DAAH01-99-C-R060 for U.S. Army Aviation and Missile Command). Fairborn, OH: Klein Associates Inc.
Malek, D. A., Alvidrez, S. A., Moon, B. M., & Wei, S. (2004). *The identification and implementation of* under Contract 1435-01-01-CT-31161 [Dept. of the Interior] for the U.S. Army Research Institute for the Behavioral and Social Sciences, *a cognitive metrics suite for the evaluation and design of tools to enhance warfighter job performance* (Final report prepared under contract number N00178-03-C-1066 for Naval Surface Warfare Center, Dahlgren, VA). Fairborn, OH: Klein Associates Inc.
Malek, D. M., Alvidrez, S., Ross, K. G., McHugh, A., Moon, B. M., Klein, G., Armstrong, A. A., & Rall, E. (2002). *High-level cognitive processes in field research* (Year One Final Report under Contract 02TA2-SP1-RT1 for U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0009). Fairborn, OH: Klein Associates Inc.
Thunholm, P. (2003). *Military decision making and planning: Towards a new prescriptive model.* Unpublished doctoral dissertation, Stockholm University, Department of Psychology, Stockholm.

Table 1

Planning

Metric	Measures
Communicate goals and intents	Successful planning depends on the planners' understanding the goals of the commissioner of the plan. Successful execution of the plan depends on those carrying out the plan understanding the thinking behind the plan.
Allow feedback between planning stages	The process of planning helps the planners learn more about the situation they are addressing and its dynamics. As a result of this learning, the planners may need to revisit earlier stages to account for their increased knowledge and understanding. Revisiting may range from elaboration of some aspect of the plan to generating a new concept for the course of action.
Manage plan risk	Potential problems in the plan are identified and tracked at every stage of the planning process. The plan is modified to account for the shortcomings, or the risk they represent is accepted. There is often a specific step in the planning process where plan execution is simulated or wargamed to identify and evaluate weaknesses in the plan.
Manage planning process	The planning process is managed so that the plan produced and the planning process are appropriate. The plan is appropriate in terms of resources available, time restrictions, and uncertainty.
Minimize handoffs and transitions	Maintaining personnel continuity through the stages of planning and into execution is the most efficient means of retaining the learning that occurs as a part of the planning process and leveraging that deep understanding in execution of the plan.

Table 2

Replanning

Metric	Measures
Notice plan is failing	Indications of plan failure can be subtle, small, and distributed, especially in the early stages of execution. Those monitoring the plan execution must be sensitized to early signals that the plan is failing.
Notice unforeseen opportunity	Focus on succeeding with the current plan should not divert from checking whether assumptions made in planning still hold and if not, whether this represents an opportunity unforeseen in planning that should now be exploited.
Lower barriers to replanning	Barriers to replanning include: fixation on existing plan that delays the start of necessary replanning; fear of unintended consequences; excessive workload; insufficient time for evaluation/wargaming; confusion about chain of authority for replanning; failure to see the significance of newly received data; difficulty in modifying plan that has many interdependencies that preclude localized changes; sunk cost effect of staying with plan because of what is already invested.
Use existing information effectively	Much data will have been collected in preparing the original plan and in monitoring its execution to this point. The relevant parts of that information collection are used in creating the new plan, and the irrelevant parts are ignored. Information is reused and reinterpreted in light of the new goals set for the new plan.
Gather appropriate information	New information will be needed to support the new plan. Planners creating the new plan must reconsider what information they might gather that will support the creation of the new plan and monitor its execution. Given that the goals have changed from the original plan, the planners must look afresh at the information needs and how they could be met.
Communicate what is changed in new plan and what is not	Those who make use of the new plan will be familiar with the old plan and be pressed for time. Execution of the new plan will be most efficient and effective if those carrying it out can easily determine what is different from the plan they are currently carrying out, and what motivated the changes.
Disseminate new plan	The plan must be disseminated to all those who received the original plan. This may be problematic because some of those will be involved in the execution of the old plan, so the new plan must be successfully communicated to them. Partial communication of the new plan may have disastrous consequences.
Monitor execution closely	In replanning situations, the planners do not have the same time to explore and understand all the consequences, side effects, and interactions inherent in the plan. When executing a replanned plan, the planners need to monitor execution of the new plan closely so that they can help respond to circumstances not considered in the replanning.