Objective measures of complexity for dynamic decision making in an interactive learning environment

Sylvain Pronovost, Jean-François Gagnon, Daniel Lafond, Sébastien Tremblay

Co-DOT laboratory, School of Psychology, 2325, rue des Bibliothèques, Québec, QC G1K 7P4 Canada
Thales Research & Technology Canada, 1405, Parc Technologique, Québec, QC G1P 4P5 Canada

sylvain.pronovost@psy.ulaval.ca, jean-francois.gagnon.22@ulaval.ca,
Daniel.Lafond@ca.thalesgroup.com, Sebastien.Tremblay@psy.ulaval.ca

ABSTRACT

Dynamic decision making (DDM)[1], is a research programme concerned with decision problems in dynamic environments, the variables of which evolve non-linearly over time, and fraught with uncertainty. Based on an interdisciplinary methodology combining a computational theoretical approach with an experimental design from applied cognitive psychology, the present work proposes objective measures of complexity for DDM. A set of seven measures relying on cognitive factors, information characteristics, and contextual constraints are suggested. Such objective measures of complexity would provide the means to (i) assess the commensurability of complex decision problems, and (ii) calibrate experiments in order to control the level of complexity of the scenarios.

Keywords: cognitive complexity; cognitive functional size; cognitive information complexity; cognitive weight complexity; complex decision making; complexity measures; cyclomatic complexity; dynamic decision making; Halstead's software metrics; interface complexity.

INTRODUCTION

Managing complex sociotechnical systems in various domains such as healthcare, defense and security, transportation and critical infrastructures, all depend on an understanding of the dynamic interrelations of such systems' components, their evolution over time, and the degree of uncertainty to which decision makers are exposed. Beyond the necessity of understanding the problem spaces they have to deal with, decision makers must also prove skillful and adaptive so as to determine how to successfully influence a complex situation, anticipate the consequences, react to surprises, and meet durable objectives. Shortcomings in the comprehension of the impacts of minute interventions, or even of long term strategies, may lead to disastrous consequences. It is thus essential to insure that decision makers are made aware of the challenges inherent to dynamic decision making, through appropriate training and supported by adequate technologies and organisational strategies.

Interactive learning environments (ILE) based on simulations of complex systems and framed as serious game can be valuable tools to facilitate performance and learning in dynamic decision making (Karakul & Qudrat-Ullah, 2008). Those simulations allow the compression of time and space in which complex decision problems unfold, and thus provide an opportunity for participants to learn about the intrinsic dynamic properties of complex systems from both a piecemeal and a holist points of view, by

[1] Also known as complex decision making (CDM).

feedback on common non-adaptive heuristics provided throughout a simulation run. The breadth and depth of feedback from such simulated environments may support a metacognitive function, i.e., support thinking about one's own decisions, and thus provide solid foundations for the development of *systems thinking,* i.e., understanding how variables interrelate when dealing with complex systems (Vester, 2007).

Objectives

The **Co**mplex **De**cision **M**aking Experimental Platform (CODEM) is a simulator and testbed used to study the impact of ILEs on the comprehension of individuals involved in complex decisions problems. More precisely, CODEM aims to study the requirements for cognitive support, *viz.* to adequately characterize what makes complexity so difficult to tackle, as well as learning to avoid non-adaptive behaviours which lead to dire consequences. This research involves human-in-the-loop experimentation on the CODEM simulation environment in order to systematically manipulate the variables and other attributes of "gamified" scenarios representing complex decision problems. The results of the CODEM project will provide substantial quantitative data to produce a model of the cognitive and behavioural determinants of the successes and failures of decision makers in complex and dynamic environments.

This paper focuses on a particular characteristic of the study of the comprehension of dynamic decision making, namely the development of an objective measure of complexity, based on an interdisciplinary methodology combining a computational theoretical approach and a design from applied cognitive psychology. This objective measure of complexity would provide the means to (i) assess the psychological commensurability of complex decision problems in DDM simulations, and (ii) calibrate future DDM experiments in order to establish a progressive scale of scenario difficulty.

Problem

Making decisions in complex situations depends on a great number of factors, but the principal characteristics of what makes decisions complex are generally framed as follows. Decision problems and decision environments, understood as complex systems, exhibit:

- uncertainty, in terms of information opacity or completeness, influencing and limiting the decision maker in its choices and in its comprehension of the system it is attempting to influence;
- non-linearities and delays in the relationships between parameters and variables, impairing the understanding of complex systems, ;
- variability/evolution over time, i.e., the dynamics of a system make it so that its parameters and variables may change states over time, both independently and as a result of human intervention.

To tackle the numerous issues related to the comprehension and control of complex, dynamic systems, our experimental testbed must represent all such features in a well-parameterized design and produce robust measures yielding some insights into participants' situational awareness and performance in attempting to stabilize a complex decision problem space. The CODEM project focuses on performance measures (goal attainment metric based on goal distance), as well as anticipation measures (difficulty measure based on prediction of future system states), for decision makers in simulated scenarios representing complex decision problems. It uses operationalized factors used as independent variables for quantitative analyses, such as the adoption or avoidance of behaviours assessed to be non-adaptive for decision problems from a systems thinking perspective (Dörner, 1989, Sterman, 1989, Funke, 1995, Bakken, 2008, Gonzalez, 2012). This paper presents a methodology, as well as some preliminary results, to support the analysis of human performance in dynamic decision making, and provide some insights in the issue of the comprehension of complex systems, by means of objective metrics and models of complexity. A formal and operationalized concept of complexity for dynamic decision problems in psychology could potentially be generalizable to other areas of research in the cognitive and behavioural sciences.

METHODOLOGY

The complexity of a decision problem, with regards to the cognitive factors, is not the focal point of research from the point of view of the numerous theories and models of complexity (Diehl & Sterman, 1995, Bar-Yam, 1997, Kinsner, 2010). According to Bar-Yam (1997), "*loosely speaking, the complexity of a system is the amount of information needed in order to describe it. The complexity depends on the level of detail required in the description.*"

A literature review of the various models and measures of complexity suggests that *algorithmic computational complexity*, as well as *structural systemic complexity*, are more appropriate for this type of problems, given that they are sensitive to information about both quantities of information, as well as the interrelatedness of components (Sipser, 1996, Kinsner, 2010).

Complexity Metrics

Seven measures have been retained to assess their relevance and their efficacy in the context of behavioural and cognitive science research, following on the positive empirical results of a number of researchers with regards to the possibility and practicality of deploying objective measures of complexity in various domains (De Silva & Kodagoda, 2013, De Silva, Kodagoda, & Perera, 2012, De Silva, Weerawarna, Kuruppu., Ellepola, & Kodagoda, 2013, Kinsner, 2010). Those objective measures of complexity are presented summarily below. Table 1 presents a synthesized view of the features and means of computation for the various models of complexity.

Cyclomatic Complexity (CC) McCabe's Cyclomatic Complexity (1976), is a measure inspired by algorithmic graph theory, which determines the number of independent paths in a control flow graph. It is computed as a simple subtraction of the number of graph nodes from graph edges, plus the number of connected components.

Halstead's Software Metrics (HM) We use a modified version of Halstead's Software Metrics (1977), an implementation-independent complexity measure of algorithms. Halstead's metric depends on the number of unique and total operators and operands, and is thus a measure of algorithmic size, a strictly informational measure of complexity**.**

Interface Complexity (IC) Cardoso's Interface Complexity (Cardoso et al, 2006), adapted from Henry and Kafura's Information Flow (1981), determines the total complexity of an algorithm as the product of structural complexity and information flow, i.e., inputs and outputs**.**

Cognitive Functional Size (CFS) Wang's Cognitive Functional Size (Shao and Wang, 2003) computes the complexity of an algorithm as a function of the product of summed inputs and outputs by the "total cognitive weight", this latter construct being itself an additive function of "basic control structures" which bear various cognitive weights determined via empirical studies**.**

Cognitive Weight Complexity Measure CWCM) Misra's Cognitive Weight Complexity Measure (2006), based on Wang's CFS, is a metric considering exclusively the total of cognitive weights, as a function of executed instructions in an algorithm**.**

Cognitive Information Complexity Measure (CICM) Kushwaha and Misra's Cognitive Information Complexity Measure (2006), also based on Wang's CFS, is a metric combining a weighted information score multiplied by the total cognitive weight.

Cognitive Complexity (CcS) Wang's Cognitive Complexity (Wang, 2007, 2009) is a more objective and rigorous measure of a system's complexity and size, because it represents its real *semantic complexity* (as opposed to mere symbolic quantification) by integrating both the *operational complexity* and the *architectural complexity* of a system in a coherent measure.

The foundational work of Wang (Shao & Wang, 2003, Wang, 2007, 2009, Wang, Kinsner, Anderson, et al, 2009) in the innovative domain of *cognitive informatics* provides some insights into more rigorous measures of complexity that include a cognitive component. By including cognitive weights in complexity measures, Wang et al. model the additional *semantic* properties of the comprehension of complex systems, beyond mere structure and information flow. Cognitive complexity is measured as the product of architectural complexity and operational complexity, where architecture refers to structural factors and operations refer to operator-system interactions (Wang, 2007, 2009).

These objective measures of complexity are tested against human behavioural data obtained through CODEM experimentations. The objective is to evaluate if these measures can account for human behaviour, comprehension and performance in scenarios of varied complexity.

Table 1. Characterization of the seven objective measures of complexity based on features included and type of calculation.

Models[2]	**Features**			**Calculation**		
	Structure	**Information**	**Cognition**	**Absolute**	**Relative (weighted)**	**Relative (scaled)**
CC	✔			✔		
HM		✔		✔		
CWCM			✔		✔	
IC	✔	✔		✔		
CFS	✔	✔	✔	✔	✔	
CICM	✔	✔	✔	✔	✔	
CcS	✔	✔	✔	✔	✔	✔

[2] CC: Cyclomatic Complexity; HM: Halstead's Software Metrics, CWCM: Cognitive Weight Complexity Measure; IC: Interface Complexity; CFS: Cognitive Functional Size; CICM: Cognitive Information Complexity Measure; CcS: Cognitive Complexity.

Apparatus

The results from previous CODEM experiments were used (Lafond, DuCharme, Gagnon, & Tremblay, 2012). This includes three scenarios, two of which were implemented in CODEM itself, *SpaceLab* and *COIN* (counterinsurgency), as well as a scenario named *Cybernetia* simulated in the Ecopolicy tm serious game (MCB Publishing House). The scenarios are contextualized implementations of dynamic decision making problems, where a participant may observe and influence the changes in variable values of key parameters in a complex system. Some of the variables must be minimized or maximized, or attain a particular threshold value, in order to achieve success (or merely avoid catastrophic failure). CODEM and Ecopolicy are turn-based computer simulations, and require that participants expend some "action points" in order to effect some changes in the system variables.

The CODEM scenarios (*SpaceLab*, *COIN*) involved 60 participants on a repeated measures design, whilst the *Cybernetia* scenario in Ecopolicy measured the performances of 30 participants, assigned to two experimental conditions (the second condition of which has been excluded for the purposes of this analysis).

The performance measures for all three scenarios, in terms of goal attainment following a computation of the goal distance between initial values and values which warrant the successful completion of a scenario, were weighted over the number of turns required to complete a given scenario, as the three scenarios did not share identical characteristics in order to establish comparable scores. Since the performance measurements of different scenarios may not warrant comparable values, given the arbitrary semantics of success or failure relative to the various variables states, their number, relationships, and the types of interventions available to influence the system, another dependent variable was included in the meta-analysis.

This variable is a set of scores for predictions of future system states, weighted and normalized in relation to actual future system states. This variable, the *anticipation* measure, is calculated through a proportion of variance accounted for (PVAF), or coefficient of determination, by comparing predictions and actual future system states, for four "turns" of the scenario simulations. It is expected that using such standardized scores may yield more objectively accurate measures of scenario difficulty, permitting comparisons of results for the different scenarios. The measures of complexity used herein may not be sensitive enough, or even capture at all, considerations of scenario difficulty, and additional features of DDM may be required in order to faithfully represent the complexity of complex decision making (see the discussion section below for further elaboration on this topic).

The following section presents the performance and the anticipation scores as measured against the seven objective measures of complexity described in the methodology section above.

RESULTS

Two types of statistical measures and two types of information visualization techniques were used to examine relationships between performance and prediction scores on the one hand, and objective measures of complexity on the other hand, as well as between the objective measures of complexity in relation to one another:

- associative tests, in the form of both the Pearson product-moment correlation coefficient (r), and Spearman's rank correlation coefficient (ρ)
- goodness of fit tests, in the form of coefficients of determination (R^2),
- a subset of the scatterplot charts for the three scenario performance and prediction scores according to the objective measures of complexity,
- a parallel coordinates chart to compare the objective measures of complexity between themselves.

Tables 2 and 3 below present the linear dependence results and the associated goodness of fit measures for the seven objective measures of complexity, with scenario performance and scenario prediction scores as dependent variables. For scenario performances, the best scores for linear dependence and goodness of fit are accounted for by the Interface Complexity model (r = -.468, R^2 = .219, and adjusted R^2 = .213), Halstead's Software Metrics (r = -.462, R^2 = .214, and adjusted R^2 = .208), and the Cognitive Functional Size model (r = -.448, R^2 = .201, and adjusted R^2 = .195), respectively. Variance in performance is thus well accounted for by the objective models of complexity, as they all range between 15.8% and 21.3%.

For scenario anticipation, linear dependence and goodness of fit in prediction scores are again better for the Interface Complexity model (r = -.407, R^2 = .166, and adjusted R^2 = .159), but in a tie with Halstead's Software Metrics (identical score), closely followed by the Cognitive Functional Size model (r = -.406, R^2 = .165, and an adjusted R^2 = .159, which ties with the above-enumerated adjusted R^2), respectively. Again, variance in anticipation is reasonably accounted for by the measures of scenario complexity, ranging between 15.1% and 15.9%.

The Spearman rank correlation coefficients, or Spearman's *rho*, tell a different story for both performance and anticipation scores. The values are ordered in a dichotomous fashion, with one model standing out – the Interface Complexity model, at ρ = -.444 for performance and ρ = -.328 for anticipation, whilst all other models fall behind at the same scores of ρ = -.403 and ρ = -.323 for performance and anticipation, respectively. This might indicate that while the IC model is more sensitive to the differences in the dependent variables, the overall majority of the objective models of complexity are indiscriminate in their ability to predict variance in the target variables. A simple explanation would be that the models of complexity are highly dependent between themselves, as they all inherit a number of features from their chronological predecessors, such as McCabe's model (for all models except HM and the CWCM), and the CFS (for CWCM, CICM, and CcS), thereby falling all along the same rank order for correlational purposes.

The four scatterplots that follow (Figures 1 to 4) present a subset of those performance and anticipation scores for the three scenarios, according to the lowest and the highest fit with the objective measures of complexity (the CWCM model and the IC model, respectively). There is a clear reversal of trend in the relationship between the complexity measures and the participants' performances and predictions, for two of the objective measures of complexity, which can be observed in Figures 3 and 4 for the Interface Complexity model, and also occurs in the case of Halstead's Metrics measure. The correlation scores and associated Figures suggest that for those two models of complexity, the differences in complexity for *SpaceLab* and *Ecopolicy* are much smaller than for the other five objective measures of complexity.

Finally, a parallel coordinates chart (Figure 5) was computed in order to represent the linear trends (or lack thereof) for the seven objective measures of complexity over the different scenarios. This information visualization technique relates the scales of the objective measures of complexity on the vertical axes, to represent linear trends and other relationships in a multidimensional space. The aforementioned divergences in linear trend between the Halstead's Metrics model and the Interface Complexity model, on the one hand, and the other five models of complexity, on the other hand, is evident in the parallel coordinates chart, but only for the *Ecopolicy* scenario.

Table 2. Monotonic and linear dependence, goodness of fit between scenario performances and objective measures of complexity[3].

	Spearman's *rho*	**Pearson's *r***	r^2	**adjusted** r^2
CC	-.403	-.427	.182	.176
HM	-.403	-.462	.214	.208
CWCM	-.403	-.402	.161	.155
IC	-.444	-.468	.219	.213
CFS	-.403	-.448	.201	.195
CICM	-.403	-.437	.191	.184
CcS	-.403	-.405	.164	.158

Table 3. Monotonic and linear dependence, goodness of fit between anticipation measures and objective measures of complexity[4].

	Spearman's *rho*	**Pearson's *r***	r^2	**adjusted** r^2
CC	-.323	-.403	.162	.156
HM	-.323	-.407	.166	.159
CWCM	-.323	-.396	.157	.151
IC	-.328	-.407	.166	.159
CFS	-.323	-.406	.165	.159
CICM	-.323	-.404	.164	.157
CcS	-.323	-.397	.158	.151

[3] All Spearman's ρ and Pearson's r values are significant at the 0.01 level, 1-and 2-tailed.

[4] All Spearmans' ρ and Pearson's r values are significant at the 0.01 level, 1-and 2-tailed.

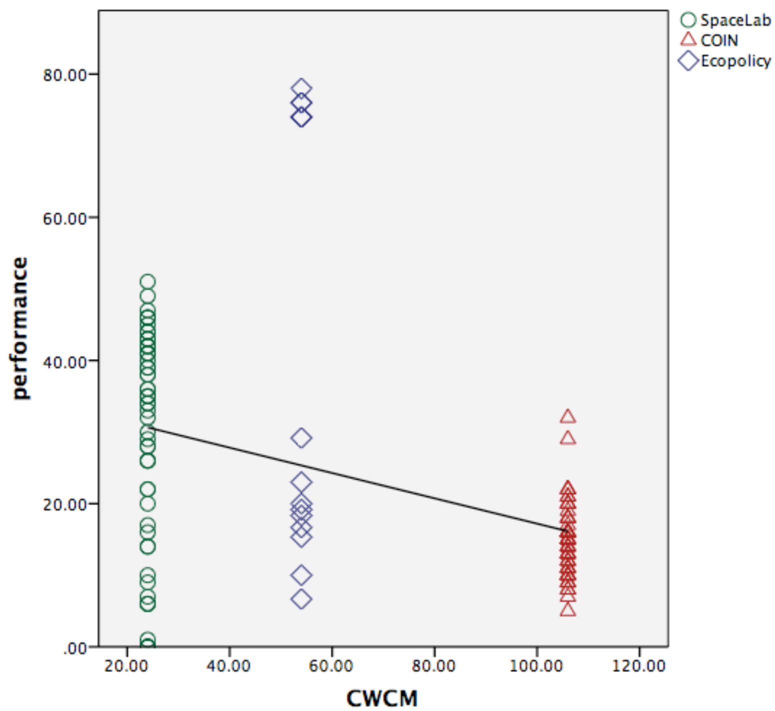

Figure 1. Scenario performances according to the Cognitive Weight Complexity Measure.

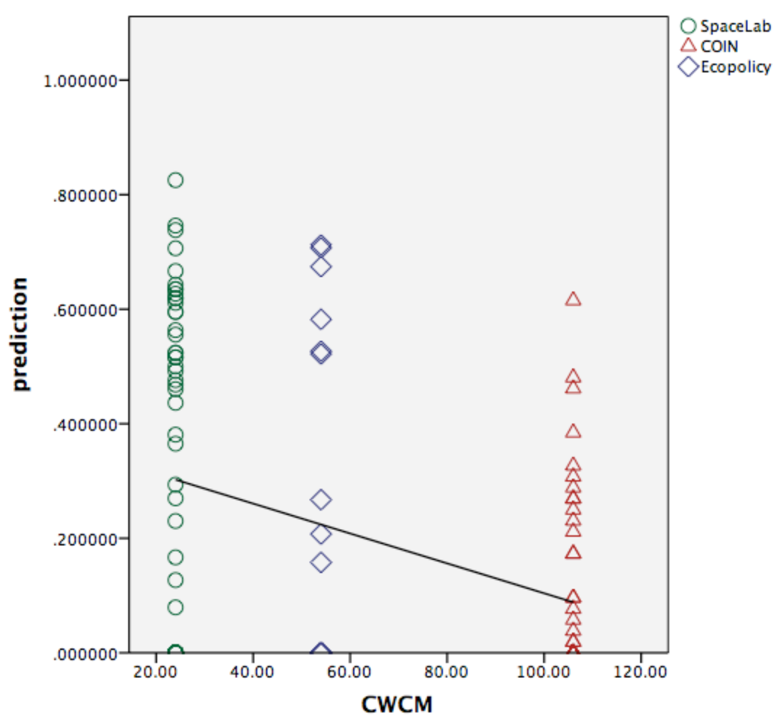

Figure 2. Prediction scores according to the Cognitive Weight Complexity Measure.

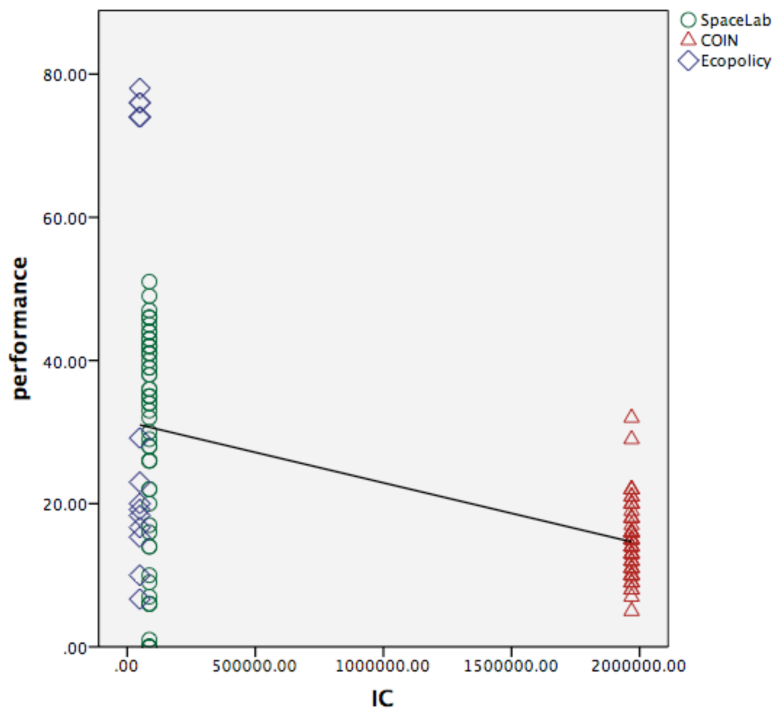

Figure 3. Scenario performances according to the Interface Complexity measure.

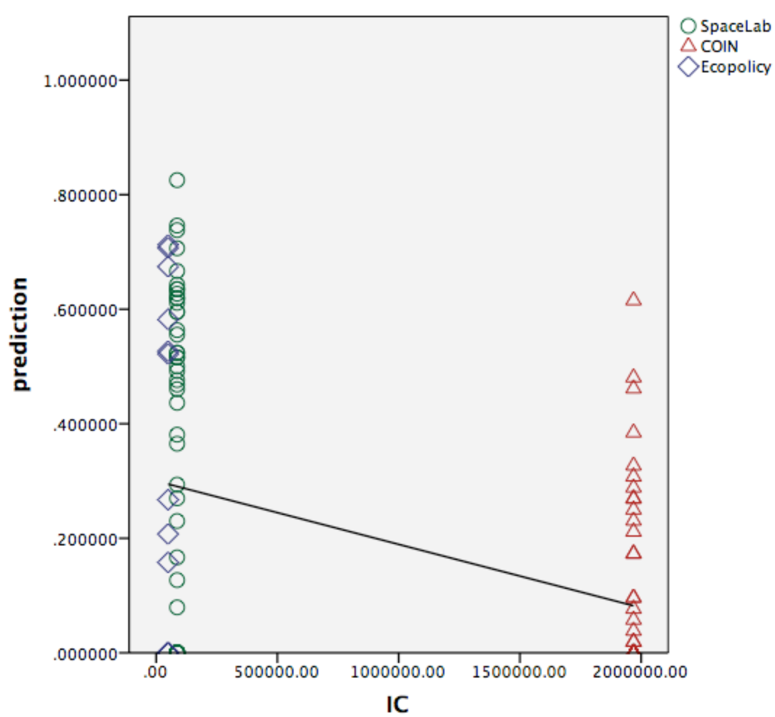

Figure 4. Prediction scores according to the Interface Complexity measure.

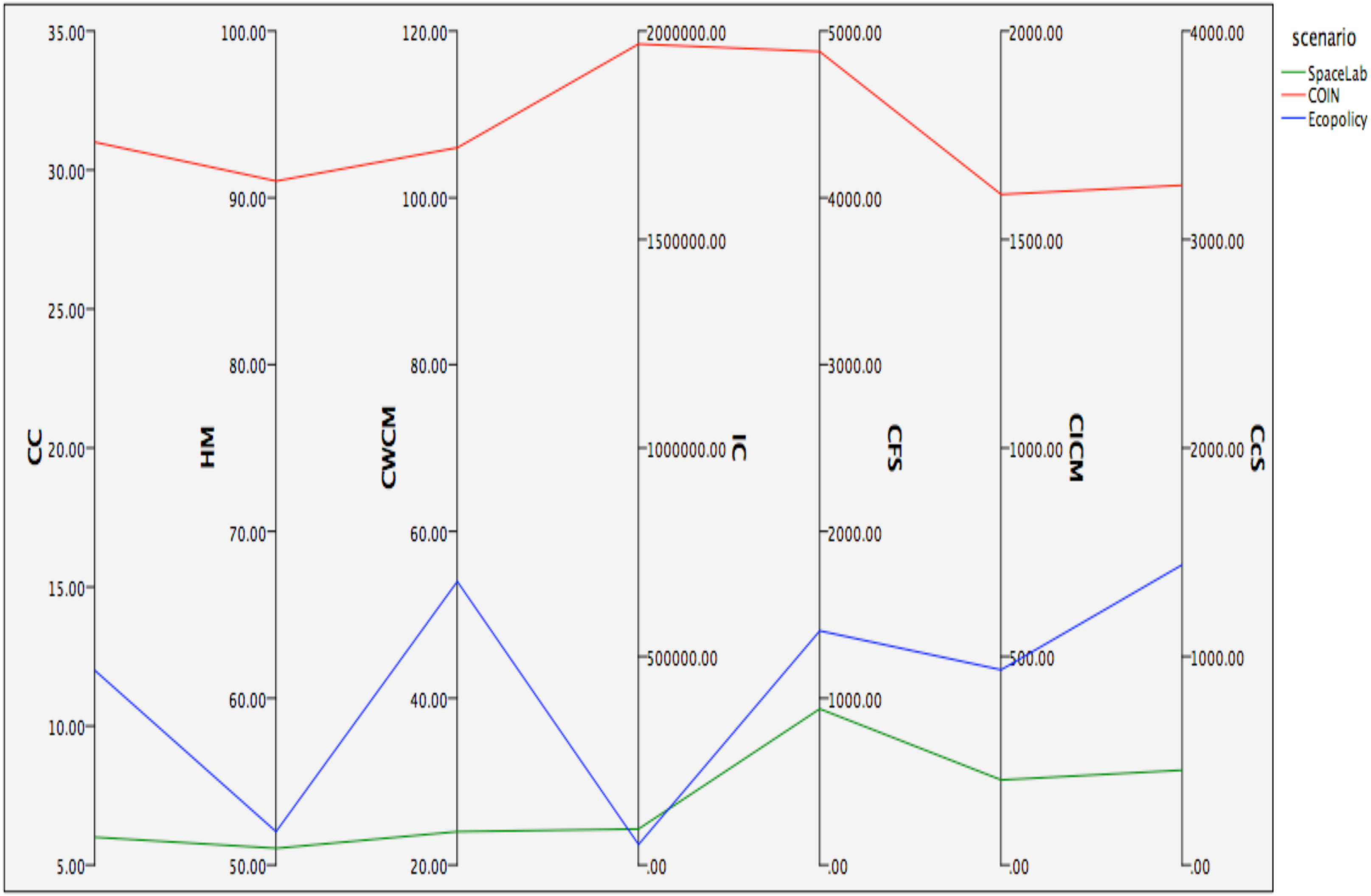

Figure 5. Parallel coordinates chart for seven objective measures of complexity across scenarios. The vertical axis represents relative (standardized) values for each measure of complexity.

DISCUSSION

The seven objective measures of complexity adapted for DDM problems have yielded significant results with regards to their explained variance on performance and anticipation measures which, while within reasonably close scores from one to another, are consistently highest for the measures most sensitive to the information flow and cognitive weights. The following sections summarize our findings and interpretations, known limitations, and future research ideas.

Dependence and Goodness of Fit

The results used in the present paper are part of a meta-analysis computed retroactively on data from dynamic decision making studies involving interactive learning environments (CODEM, Ecopolicy). While the number of participants was sufficient to explore linear dependencies and goodness of fit for objective measures of complexity between themselves and with human performance in various DDM scenarios, additional data is required to validate any claim to the linear dependence between such models of complexity and performance in DDM.

The current study used three different scenarios of varying complexity and human performance scores, and the results are so far counterintuitive with regards to some of the findings from De Silva et al (De Silva & Kodagoda, 2013, De Silva, Kodagoda, & Perera, 2012, De Silva, Weerawarna, Kuruppu, Ellepola, & Kodagoda, 2013) and from Wang (Shao & Wang, 2003, Wang, 2007, 2009). As can be observed from Table 1, the highest correlation and regression scores between an objective measure of complexity and human performance in the three scenarios come from the Interface Complexity model, Halstead's Software Metrics, and the Cognitive Functional Size model, respectively. The caveat is that both IC and HM are the complexity measures that are also the two with the least linear dependence with other objective measures of complexity (Figure 5).

De Silva et al (De Silva, Kodagoda, & Perera, 2012, De Silva, Weerawarna, Kuruppu, Ellepola, & Kodagoda, 2013) had very high correlation scores (using Spearman's rank correlation coefficient) for monotonic dependence between *subjective* measures of complexity and objective models such as Shao and Wang's Cognitive Functional Size (.863 and .903 in the 2012 and 2013 articles, respectively), and McCabe's Cyclomatic Complexity (.752 in the 2012 article), with lower Spearman's *rho* values for the CICM and CWCM measures. We conjecture that this could be explained by one or more of the three following factors:

- subjective measures of complexity may yield higher linear dependence with objective measures of complexity than performance data in a complex decision problem,
- the assessment of the complexity of algorithms may be incommensurable with, or substantially different from, the assessment of dynamic decision problems (including different sensitivities to the various structural, informational, and cognitive features elicited in the measures of complexity),
- the information contained in the human performance data, and/or the information about the intrinsic properties of the three DDM scenarios themselves, is not heterogeneous enough to yield higher discrepancies in correlation and regression values, compared to the ten algorithms used in De Silva et al's three papers, for which the subjective and objective measures of complexity were analyzed.

The discrepancies between Pearson product-moment correlation coefficients and Spearman rank correlation coefficients in our results may be attributable to violations to the ideal conditions for linear correlation. While the Pearson r values are higher and diverse, suggesting higher model sensitivity and higher dependence, a number of essential assumptions to the robustness of Pearson's correlation undermine the accuracy of our data. Firstly, there are a few outliers in the Ecopolicy dataset, but the *n* is already pretty low, given our restriction to the conditions of this experiment that we have opted to compare with the CODEM scenarios, so we did not remove those outliers for the sake of consistency. Secondly, the performance and anticipation scores are skewed in distribution. The presence of outliers and skewed variables sometimes favors higher Pearson correlation scores than is warranted. In future analyses, a nonlinear modeling approach will be explored, by analyzing DDM data via the proposed *general monotone model* (GeMM) of Dougherty and Thomas, (2012) which is more appropriate for monotonic and nonlinear relationships.

Feature-Dependent Metrics of Complexity

The seven measures of complexity used in this paper span models and theories of complexity pertaining to structural, informational, and cognitive features of complex systems. Some do in a concatenative fashion, i.e., by conservative extension to another metric, such as Cardoso's Interface Complexity which adds information flow to a structural component akin to McCabe's Cyclomatic Complexity, while Misra's Cognitive Weight Complexity Measure and Kushwaha and Misra's Cognitive Information Complexity Measure both amend Shao and Wang's Cognitive Functional Size metric.

The selective or concatenative emphasis on structural, informational, and/or cognitive factors may be preponderant in certain domains of complexity theory, such as algorithmic complexity theory and software engineering, as studied by De Silva et al (De Silva & Kodagoda, 2013, De Silva, Kodagoda, & Perera, 2012, De Silva, Weerawarna, Kuruppu, Ellepola, & Kodagoda, 2013). Those researchers had a very clear focus in mind: the effectiveness and practicality of adapting such measures of complexity to software engineering best practices. Our focus is both fundamental and experimental: which of the objective complexity measures, amongst the myriad available, can leverage our understanding of complex/dynamic decision problems? Further experimentation may reveal fundamental trends or specific requirements of complexity modeling for behavioural and cognitive science research endeavours.

Sensitivity to Absolute and Weighted Features

The structural, informational, and cognitive factors involved in the above mentioned objective measures of complexity are not only selective in what they measure, they also adopt significantly different scales and calculations for the same or similar factors. Wang's Cognitive Complexity, for instance, modifies the calculations of the Cognitive Functional Size so as to weight the informational component of complexity with a ceiling value (3 inputs or more is treated as a value of 3, similarly for outputs), while Kushwaha and Misra's Cognitive Information Complexity Measure weights the informational component as a ratio of the sum of operators and operands to the quantified symbolic information (effective lines of code, *eLOC*). Similarly, what counts as structural complexity may be an absolute value as in the case of McCabe's Cyclomatic Complexity and Cardoso's Interface Complexity, or a weighted feature directly included in the cognitive component of complexity, such as is the case for the CFS, CWCM, and CICM metrics. Further, it may be that the relative contribution and weight of a feature ends up being revised on empirical grounds, such as the cognitive component of complexity in Wang's Cognitive Complexity, renamed the operational complexity (Wang, 2007, 2009).

Based on the observations gathered in the results section, one quirk of such computational and empirical choices is that it may be entirely possible that objective measures of complexity do not model the same scenarios in a linearly dependent trend. That is, for a given dynamic decision problem, a number of the measures of complexity depart from their linear trend because the weight of a given feature in one objective measure or another looms larger than the average. For example, Halstead's Metrics and the Interface Complexity model clearly diverge from the linear trend of other measures in the case of the Ecopolicy scenario in the parallel coordinates charts (Figure 5). This is because HM and IC are very sensitive to the informational component of complexity. HM computes the sum total of operators and operands, including inputs and outputs, while IC uses an absolute computation of the squared sum of inputs and outputs, both of which are very low in the Ecopolicy scenario, relative to the CODEM scenarios.

Known Limitations

A number of constraints affect the current formulation of the objective models of complexity, based on arbitrary choices in the experimental design, on computational capabilities and means, as well as on some fundamental principles intrinsic to the models used.

The first issue is that of *static* versus *dynamic* measures of complexity. This imposes a constraint to our modeling endeavour pertaining to a difference between *structural* and *execution* semantics in the description of a complex system. Our objective measures of complexity do not capture variability in inputs and outputs for every turn of a simulated scenario execution, nor do they capture the number of interventions used. The number of functional, i.e., relevant, parts in a complex system varies over time (Brehmer & Allard, 1991), and the intrinsic dynamics of a system may impact its own complexity (Kerstholt & Raaijmakers, 1997). The models of complexity focus only on the initial conditions of a DDM scenario, i.e., they capture a snapshot of the initial range of possible inputs and outputs, as well as the maximum number of possible interventions that can be effected over the system. A possible venue to capture the execution semantics would be to apply a *dynamic complexity* measure which computes structural complexity features across a number of dimensions and over time, such as the *variance fractal dimension trajectory* (Kinsner, 1994).

Secondly, the measures of complexity used herein are parameterized and implemented differently than their algorithmic and structural models of origin, because there exists both intrinsic and practical differences between algorithms and decision problems. The scenarios in eCODEM are built around constructs such as variables, actions, influences between variables, action points and output functions as information flow, etc., and the operationalization of the objective models of complexity require some extensive tweaks in repurposing and adapting their metrics to a substantially different domain. For some complexity metrics, absolute measures are used in total or for subcomponents, while in other measures, weighted scores are used. Only one objective model of complexity, i.e., Wang's Cognitive Complexity (2007, 2009) addresses this issue in a way that translates well from algorithmic complexity analysis to decision analysis. Wang proposes the concept of *relational complexity*, $C_r(S)$, as the maximum potential or the upper limit of the operational complexity of a given system, with the symbolic complexity being its lower boundary. With well-formulated and operationalized concepts for *symbolic*, *architectural*, *operational*, *cognitive*, and *relational* complexities, the translation of dynamic decision making problems into effectively computable complex systems is made possible.

Finally, as mentioned in the methodology section, the objective measures of complexity used in this research may not sensitive to scenario *difficulty*, *viz*. the goal distance between the system's initial variable states and a system's set of variable states that represents the successful completion of a DDM scenario. Indeed, the features of a complex system, such as structural intricacies or information flow volume, neither translate directly, nor in isolation from one another, into some measure of cognitive demand on a decision maker (Gonzalez, Vanyukov, & Martin, 2005). A complex system with twice as many interrelated components is not twice as difficult to understand or influence (Funke, 1988). Gonzalez et al (2005) also note that "*decision makers may be able to achieve the goals of some DDM systems by considering only a subset of system variables*". This is consistent with Mackinnon and Wearing's (1980, also mentioned in Karakul & Qudrat-Ullah, 2008) finding that interactions between subsystems in DDM facilitate overall performance. In order to insure that a complex system does not yield to such narrower strategies, we must increase the interrelatedness of components so as to eliminate such subsystems. As the authors note, "*in such tightly coupled systems, complexity remains relatively constant because each dynamic decision requires consideration of most or all system variables*". Future models of complexity will factor an objective measure of difficulty, in terms of goal-distance of initial decision space values from values warranting a successful scenario, and this will be accomplished in a twofold methodology: firstly, an objective measure of difficulty will be used as a an independent predictor of dependent measures, and secondly, all seven objective models of complexity described herein will incorporate a component score factoring the same computation for difficulty, in order to observe whether difficulty measures are a viable feature for complexity analysis in dynamic decision making research on their own, or compounded with other factors of complexity (structure, information flow, cognitive weights).

Way Ahead

Additional performance data points are required to assess how robust are the scores for the linear dependence and goodness of fit of the seven objective measures of complexity used in this meta-analysis. Future experimentation will include additional dependent variables besides overall scenario performance and anticipation scores, including subjective measures of complexity, confidence measures, and performance measures on working memory and fluid intelligence (*n*-back test). The scenarios in CODEM will involve participants with and without training in systems thinking, as well as the introduction of a cognitive tutor to provide feedback on non-adaptive heuristics use. New scenarios will be used, including a different simulation platform (Democracy 2 ^{tm}) to cross-validate the findings from the CODEM platform.

There is a possibility that the effect of complexity and difficulty on DDM performance falls along a just-noticeable difference threshold as formulated by the Weber–Fechner law[5]. This would amount to a better fit if the relationship between increasing levels of complexity in decision space, on the one hand, and overall performance and/or perceived difficulty on the other hand, were modeled according to a nonlinear statistical model. Future experimentation and analyzes will take into account this differential limen to test whether DDM performance and/or difficulty do in fact follow a logarithmic scale.

---

[5] Reber, A. S. (1985) *The Penguin dictionary of psychology*. London: Penguin Books.

ACKNOWLEDGEMENTS

This research was funded by the Department of National Defence and the Natural Sciences and Engineering Research Council. The authors would like to thank Michel B. DuCharme, from Defence R&D Canada – Valcartier, for ongoing collaborations on complex decision making research through the DRDC Applied Research Project # 10bp, and Marie-Ève St-Louis, from Université Laval, for her invaluable help with regards to data collection.

REFERENCES

Bakken, B. E. (2008). On improving dynamic decision-making: implications from multiple-process cognitive theory. *Systems Research and Behavioral Science, 25* (4), 493-501.

Bar-Yam, Y. (2003). *Dynamics of complex systems*. Westview Press.

Brehmer, B., & Allard, R. (1991). Dynamic decision making: The effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision making: Cognitive models of cooperative work*. Chichester: Wiley.

Cardoso, J., Mendling, J., Neumann, G., & Reijers, H. A. (2006). A discourse on complexity of process models. In *Business process management workshops, LNCS vol. 4103* (pp. 117-128). Springer Berlin Heidelberg.

De Silva, D. I., & Kodagoda, N. (2013). Applicability of Weyuker's Properties Using Three Complexity Metrics. In *proceedings of the 8th International Conference on Computer Science & Education (ICCSE 2013).*

De Silva, D. I., Kodagoda, N., & Perera, H. (2012). Applicability of three complexity metrics. In *Advances in ICT for Emerging Regions (ICTer), 2012 International Conference on* (pp. 82-88).

De Silva, D. I., Weerawarna, N., Kuruppu, K., Ellepola, N., & Kodagoda, N. (2013). Applicability of Three Cognitive Complexity Metrics. In *proceedings of the 8th International Conference on Computer Science & Education (ICCSE 2013)*.

Diehl, E., & Sterman, J. D. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes, 62*(2), 198-215.

Dörner, D. (1989). *Die Logik des Mißlingens*. Reinbek bei Hamburg: Rowohlt.

Dougherty, M. R., & Thomas, R. P. (2012). Robust decision making in a nonlinear world. *Psychological review*, *119*(2), 321.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking Reasoning, 7*(1), 69–89.

Funke, J. (1995). Experimental Research on Complex Problem Solving. In P. Frensch & J. Funke (Eds.) *Complex problem solving: The European perspective* (pp. 3–25). NJ: Lawrence Erlbaum Associates Publishers.

Funke, J. (1988). Using Simulation to Study Complex Problem Solving: A Review of Studies in the FRG. *Simulation & Gaming, 19*(3), 277–303.

Gonzalez, C. (2012). Training decisions from experience with decision making games. *Adaptive technologies for training and education*, 167-178.

Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior, 21*(2), 273–286.

Halstead, M. H. (1977). *Elements of Software Science (Operating and programming systems series)*. Elsevier Science Inc.

Henry, S., & Kafura, D. (1981). Software structure metrics based on information flow. *Software Engineering, IEEE Transactions on*, (5), 510-518.

Karakul, M., & Qudrat-Ullah, H. (2008). How to improve dynamic decision making? Practice and promise. In *Complex Decision Making*. Springer Berlin Heidelberg.

Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In Ranyard, R., Crozier, R. W., & Svenson, O. (Eds.), *Decision making: Cognitive models and explanations*. Norwood, NJ: Ablex.

Kinsner, W. (2010). System complexity and its measures: How complex is complex. In *Advances in Cognitive Informatics and Cognitive Computing* (pp. 265-295). Springer Berlin Heidelberg.

Kinsner, W. (1994). Batch and real-time computation of a fractal dimension based on variance of a time series. *Dept. of Electrical & Computer Eng., University of Manitoba, Winnipeg, Canada DEL94-6*.

Kushwaha, D.S., & Misra, A.K. (2006). A moditied cognitive information complexity measure of software. *ACM SIGSOFT Software Engineering Notes, 31* (1), 1-4.

Lafond, D., DuCharme, M. B., Gagnon, J. F., & Tremblay, S. (2012). Support requirements for cognitive readiness in complex operations. J*ournal of Cognitive Engineering and Decision Making, 6*(4), 393-426.

Mackinnon, A. J., & Wearing, A.J. (1980). Complexity and decision making. *Behavioral Science, 25*(4), 285–292.

McCabe, T. J. (1976). A complexity measure. *Software Engineering, IEEE Transactions on*, (4), 308-320.

Misra, S. (2006). Complexity measure based on cognitive weights. *International Journal of Theoretical and Applied Computer Sciences,1* (1),1-10.

Shao, J., & Wang, Y. (2003). A new measure of software complexity based on cognitive weights. *Canadian Journal of Electrical and Computer Engineering, 28*(2), 69-74.

Sipser, M. (1996), *Introduction to the Theory of Computation*. PWS Publishing Co.

Snowden, D. (2000). Cynefin: a sense of time and space, the social ecology of knowledge management. In Despres, C and Chauvel, D (Eds), *Knowledge Horizons: The Present and the Promise of Knowledge Management*, Butterworth-Heinemann: Oxford.

Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*, 301–335.

Vester, F. (2007). *The Art of interconnected thinking: Tools and concepts for a new approach to tackling complexity*. Westermann-Schulbuchverl.

Wang, Y. (2009). On the cognitive complexity of software and its quantification and formal measurement. *International Journal of Software Science and Computational Intelligence (IJSSCI), 1*(2), 31-53.

Wang, Y. (2007). The theoretical framework of cognitive informatics. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 1*(1), 1-27.

Wang, Y., Kinsner, W., Anderson, J. A., Zhang, D., Yao, Y., Sheu, P., ... & Chan, C. (2009). A doctrine of Cognitive Informatics (CI). *Fundamenta Informaticae, 90*(3), 203-228.