Review of Cognitive Metrics for C2

Mandy Natter
Jennifer Ockerman, PhD
Leigh Baumgart

APL
The Johns Hopkins University
APPLIED PHYSICS LABORATORY

Agenda

- Measures Overview: What to consider
- Summary of Cognitive C2 Metrics
  - Workload
  - Situation Awareness
  - Decision Making
  - Collaboration
- Future Work

Measures Overview: What to consider

- **Measurement types:** qualitative, quantitative, subjective, objective
- **Measurement scales:** nominal, ordinal, interval, ratio
- **Number of participants:** power analysis from pilot study
- **How measures are analyzed:** descriptive statistics or inferential statistics
- **Raters:** self reports, experimenter, subject matter experts
- **Timing:** real-time, post hoc
- **Measurement Criteria**

Measures Overview: What to consider (cont'd)

Measurement Criteria

Criteria	Description
Validity	Measuring the mental construct of interest[1]
Repeatability / reliability	Same results of the mental construct when tests are administered more than once[1]
Sensitivity	Detect changes in the level of the mental construct imposed by task difficulty or resource demand
Diagnosticity	Discriminate the amount of the mental construct imposed on different operator resources (e.g., perceptual versus processing versus motor resources)
Selectivity	Sensitive to differences only in the cognitive construct of interest (e.g. cognitive demands as opposed to physical workload or emotional stress[1])
Intrusiveness	Interfere with performance on the primary task
Implementation requirements / convenience	Ease of implementing specific assessment technique (e.g., instrumentation requirements or operator training)
Operator Acceptance	Willingness on the part of the operators to follow instructions and actually utilize a particular assessment technique

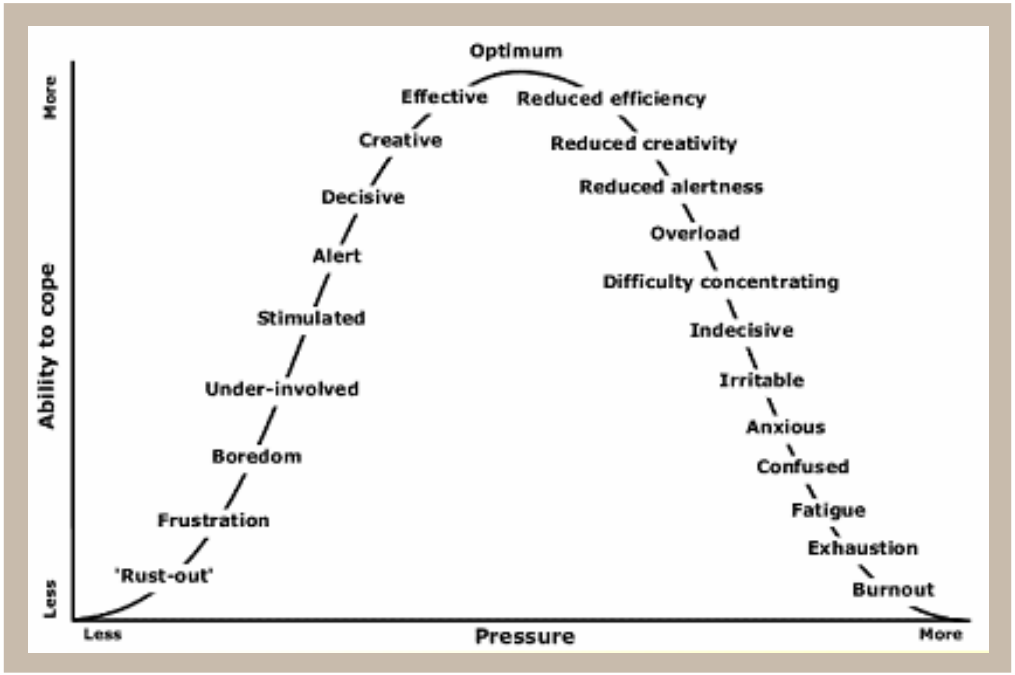
[1]Zhang & Luximon, 2005

Interactions of Cognitive C2 Metrics

Workload

- **Workload is the portion of human's limited capacity that is required to perform a particular task[1]**
  - **In C2, cognitive workload is of most interest[2]**

- **Appropriate workload is most important**
  - **Too low or too high can both be bad**



[1]O'Donnell & Eggemeier, 1986 [2]Zhang & Luximon

Workload Measurement Techniques

- **Primary-task measures: Quality (speed & accuracy) of primary task**
  - Pro(s): Objective; related to performance
  - Con(s): May reflect data or system limitations vs. human, may have low sensitivity when task is easy

- **Secondary-task measures: Quality (speed & accuracy) of secondary task**
  - Pro(s): High validity- helps predict residual resources in the event of a failure, can compare the workload of two different primary tasks
  - Con(s): Interference with primary task, must match the resource demands of the primary task

- **Subjective measures: Ratings by person doing task or observing subject matter expert**
  - Pro(s): Low cost, ease of use, general non-intrusiveness, high validity, high sensitivity (at times more than objective measures)
  - Con(s): Confounding factors, short-term memory constraints, non-continuous (do not reflect changes in workload during the task)

- **Physiological measures: Use of physiological measures to objectively assess workload**
  - Pro(s): Continuous, very sensitive, general non-intrusiveness
  - Con(s): New and immature area, confounds (e.g., individual differences, noise)

Situation Awareness (SA)

- **Situation awareness[1] is**
  - **the *perception* of elements in the environment within a volume of time and space (level 1)**
  - **the *comprehension* of their meaning (level 2)**
  - **the *projection* of their status in the near future (level 3)**

- **Military terminology[2]**
  - **Situational Awareness (level 1)**
  - **Situational Understanding (level 2)**
  - **Situational Assessment (level 3)**

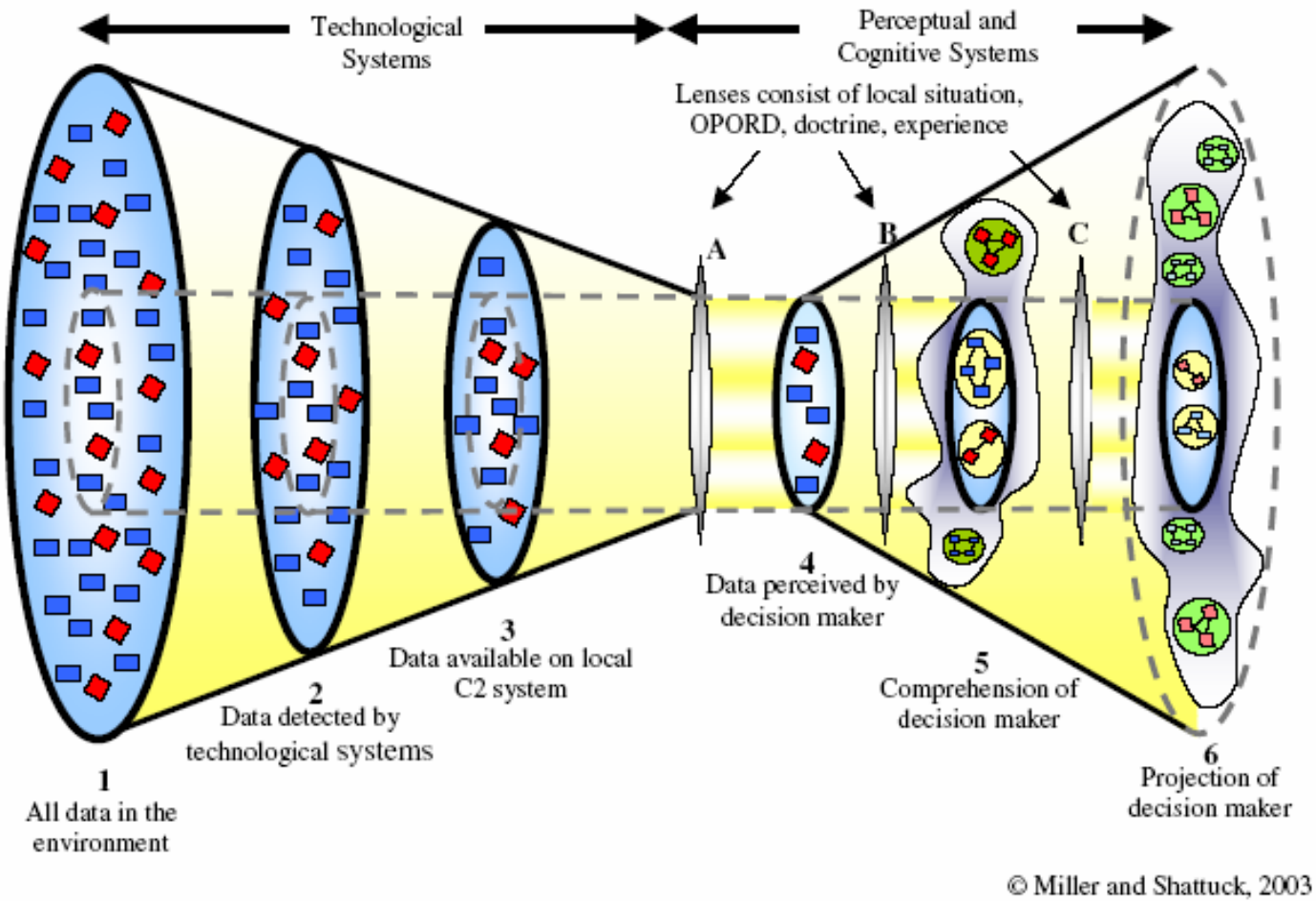
- **Sensemaking is the process to arrive at and maintain SA**

[1] Endsley, 1995 [2]Gartska & Alberts, 2004

APL

Situation Awareness (SA)

- **Technical and Cognitive components have been explored**



Cognitive Situation Awareness Measurement Techniques

- **Explicit SA measures: Ask what SA is to determine level**
  - Pro(s): Validated technique (e.g., SAGAT)
  - Con(s): Intrusive- may disrupt primary task, confound SA with probes, laborious to create probes
- **Implicit SA measures: Infer what SA is to determine level**
  - Pro(s): Easy to obtain, less intrusive than explicit measures
  - Con(s): Only simple responses and behaviors, assumption driven
- **Subjective SA measures: Ratings by person doing task or observing subject matter expert**
  - Pro(s): Does not need to be customized for different domains, easily employed, high validity
  - Con(s): May be confounded by performance and workload, usually post hoc- so rely on memory, individual differences, inter-rater reliability
- **Team SA measures: Multiple methods used**
  - Pro(s): Most C2 environments involve teams
  - Con(s): Immaturity, complexity (e.g., status within the team, lack of control, differing expectations, prevented action by other team members)

Decision Making

- Decision making is a complex process, not just the result.

- Involves selecting options from alternatives, where:
  - some information pertaining to the option is available
  - time allotted is longer than a second
  - there is uncertainty or risk associated with the selection[1]

- Information component[2]
  - provides “right” information, to the “right” person, at the “right” time
  - determine using cognitive engineering knowledge elicitation techniques (e.g., task analysis, cognitive task analysis, & cognitive work analysis, etc.)

- Human component
  - selection of, or at least responsibility, for COA
  - rational or analytical decision making[3]
  - naturalistic or intuitive decision making[4]

[1]Wickens & Hollands, 2000 [2]Means & Burns, 2005 [3]Azuma et al., 2006 [4]e.g., Klein, Calderwood, & Macgregor, 1989

Decision Making Measurement Techniques

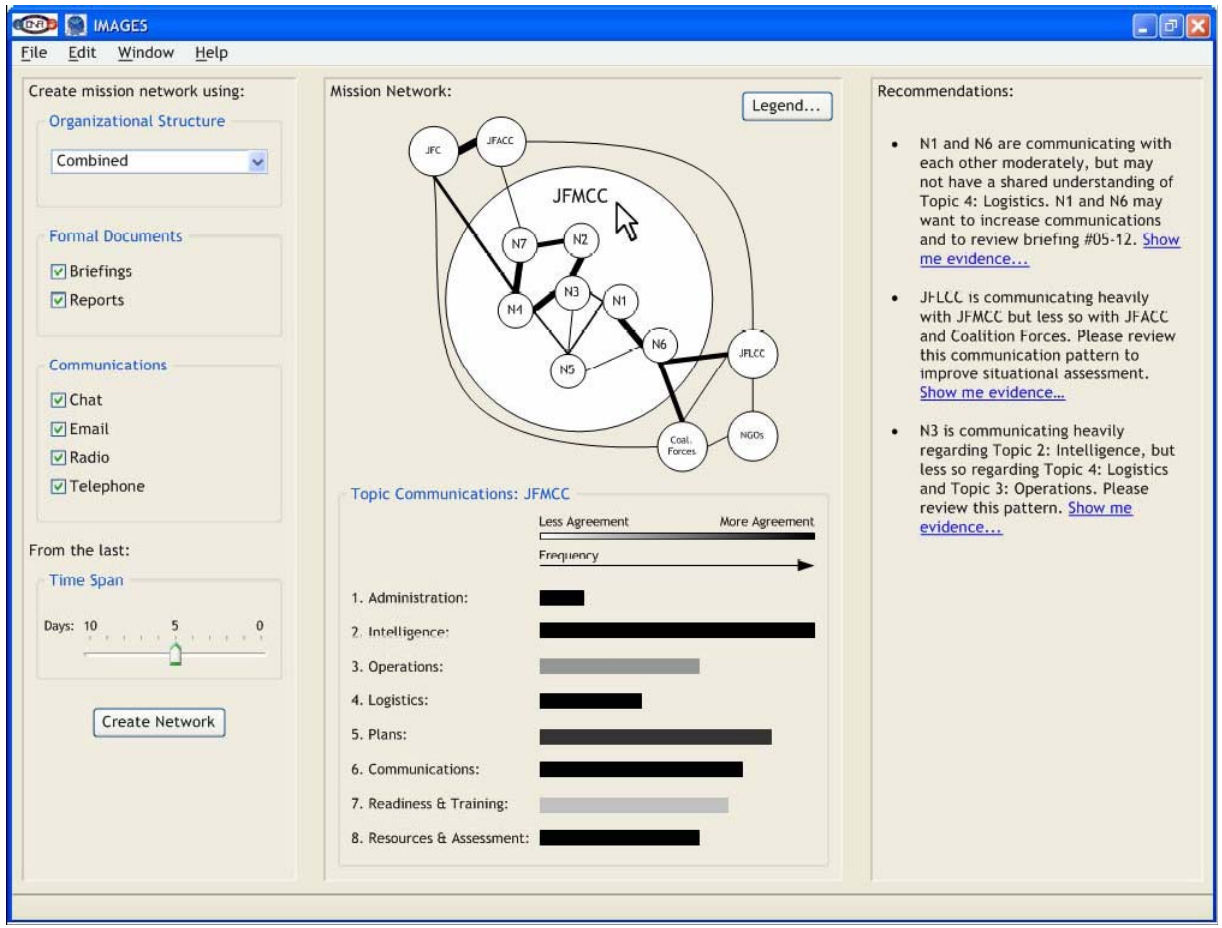
- Complicated due to difficulties
  - in defining a “good” decision
  - influence of many factors (hard to equate decision making with mission effectiveness)
  - observing or eliciting strategies
  - continuous nature of some decisions

- Result-based measures: Measure quality (accuracy and timeliness) of decision
  - Pro(s): Easy to employ, objective, observable, related to performance
  - Con(s): Doesn’t provide decision rationale, could be luck or chance, may not be “best” decision

- Process-based measures: Measure appropriateness of strategies and evaluate information used
  - Pro(s): Understand “why”, lead to improved C2 processes
  - Con(s): Some processes not observable, difficult to represent, resource intensive, difficult to assess reliability and validity

Communication and Collaboration

- **Communication is expression and may include information sharing**
  - **prerequisite for collaboration**
- **Collaboration involves leveraging the information of others to reach or meet a goal or objective**



Freeman et al., 2006

Collaboration Measurement Techniques

- **Technical-based measures: evaluate interconnectivity, modes available for communication, communication network, etc.**
  - Pro(s): Easy to automate collection, trend analysis
  - Con(s): Implicit

- **Human-based measures**
  - **time-spent collaborating, content and frequency of collaboration, use of collaboration modes, etc.**
  - **goal of collaboration**
  - **social network and knowledge distribution diagrams[1]**
  - **leadership and experience**
  - Pro(s): Understand elements of team understanding and decision making, non-intrusive
  - Con(s): Resource intensive, can be difficult to represent, difficult to assess reliability and validity

- **Content analysis is usually very time and labor intensive**

[1]**Freeman et al., 2006**

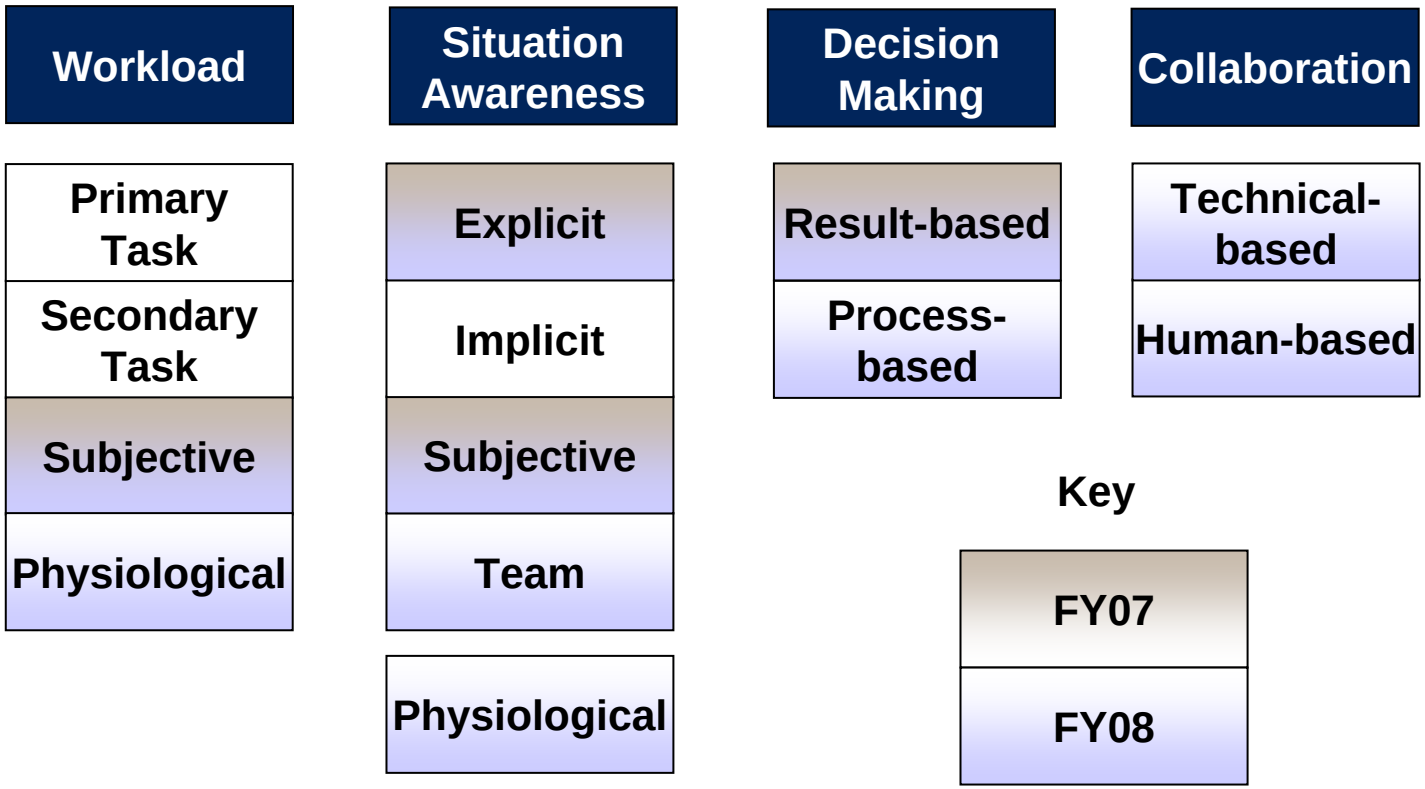
APL

Cognitive Metrics for C2 Research

- Issues
  - Still much debate within the cognitive engineering community on appropriate definitions and metrics
  - Most metrics still focused on an individual
  - Workload, SA, DM, and collaboration are highly interdependent
  - Not a lot of automation available to collect and analyze collected data

- Mitigations
  - Use suite of complimentary and overlapping measurement techniques
  - Design the evaluation and the analysis ahead of time
  - Use the automation that is available to collect and analyze data

JHU/APL Metrics Overview (FY07 and FY08)



Exploring these metrics in two studies
Physiological and Collaboration

Questions?