Maneuvering in the Social Media Space: Assessing Performance and Effectiveness

Morgan R. Borders, William A. Ross, Michael L. Williams
Cognitive Performance Group
Portsmouth, VA
{Morgan, Bill, Michael}@cognitiveperformancegroup.com

ABSTRACT

Within the information environment (IE), adversaries exploit social media using their understanding of culture, beliefs, heuristics, and biases. Many challenges must be overcome to compete with these adversaries for global information advantage, including outwitting a highly adaptive, well-trained adversary; maintaining pace with advances in data science; the ubiquity of information technologies; and understanding relevant human performance capabilities. Further, effectively planned and implemented information operations produce measurable influence on decision making, perceptions, and human behaviors, but these effects can be difficult to measure. The current research began addressing these challenges by developing a method to examine and assess human sensemaking, problem solving, and decision making in the IE that provides the necessary data to inform decision making. Developing a well-rounded, technology-agnostic, and widely applicable assessment tool involved taking a systems perspective of information operations that combines the physical, informational, and cognitive aspects of the IE. Based on a previously developed mastery model of information maneuver analysts, we created a Behaviorally Anchored Rating Scale (BARS) consisting of 59 measures of performance associated with measures of effectiveness that can be used across the information domain. We further enhanced the BARS and facilitated reliable data collection by using the Field Assessment System (FAS), a complementary ONR project, to digitalize the BARS presentation and performance ratings. Leveraging Power BI for data visualization and performance analytics across several user-relevant dimensions of performance made data synthesis user-friendly which affords easy interpretation to inform future research. This novel approach to analyst and planner performance assessment allows research across domains to focus on cognitive performance beyond machine use in the IE, which offers application in any field. In this paper, we illustrate the foundation and development of the OMEN measurement and assessment system from the perspective of human performance and not the enabling technologies.

ABOUT THE AUTHORS

Morgan R. Borders is a Scientist II at the Cognitive Performance Group with an MS in Human Factors Psychology. Her research interests include the intersection of the cognitive and social aspects of complex real-world problems, such as the use of social media in influence campaigns. Methodologically, she applies a contextualized multi-level perspective combining lab and field-based methods, such as computational language analysis.

William A. Ross is a Principal Research Scientist at the Cognitive Performance Group. His research interests include investigation into the nature and acquisition of rapid cognition skills for tactical decision makers and the application of experiential learning to support the development and assessment of expertise. Bill Ross combines extensive experience in measurement and assessment of dynamically complex systems with a broad background in operational design, planning, and analysis.

Michael L. Williams is a retired Army Information Operations officer and consultant. He entered the U.S. Army Chemical Corps in late 1985 and transitioned to Military Intelligence a few years later. He became an IO Officer in 1999 and an IO Team Leader. He later led the training and deployment of IO teams supporting tactical and theater headquarters. He finished his military service while assigned to the Joint Staff J-39. He then became an advisor and consultant to OUSD(I) and OUSD(Policy) working on strategic communication and information operations policy.

Maneuvering in the Social Media Space: Assessing Performance and Effectiveness

Morgan R. Borders, William A. Ross, Michael L. Williams
Cognitive Performance Group
Portsmouth, VA
{Morgan, Bill, Michael}@cognitiveperformancegroup.com

INTRODUCTION

Information, as a domain of warfare, is vital to every aspect of operations and mission accomplishment. Imagine a context where cognitive warfare among competing networks requires an understanding of how intelligent agents and bad actors use social media platforms to disseminate information for influencing our actions and behaviors. Project OMEN is a science and technology initiative for harnessing the power of information and network analysis to support Information Operations (IO) planning and decision making. The aim of the OMEN project is to demonstrate an operational capability that uses social media analytics and data visualization for planning, conducting, and assessing effective information maneuvers that complement other military and civil operations. OMEN tools provide a capability to recognize, detect and counter disinformation as part of operations in the information environment (IE) that is tailored to satisfy command-specific requirements. The current effort used decision research methods to create a set of technology agnostic measures for assessing individual and team performance within an IO Cell that uses the OMEN tools and methods. The next sections introduce the OMEN project, the IE, and the relevant tools and methods. The sections following the introductions describe the current effort including the development of the OMEN behaviorally anchored rating scale (O-BARS), the proof-of-concept demonstration and testing the O-BARS, and the potential applications.

OMEN Overview

At a high level, OMEN is designed to assist analysts and planners with rapid sensemaking of social media content and allow an informed judgment about the effect of that content on behavior that could affect the outcome of military operations. Further, the OMEN program is researching potential capabilities that create an ability to accurately reflect the complex social media IE in exercises and simulations. OMEN is an ecosystem comprised of data management and analysis solutions as an integrated set of tools, technologies, algorithms, procedures, measures, and information maneuvers designed to influence decision making, alter perceptions, and change behaviors, which satisfy service requirements for a capability to explain and exploit the IE.

The OMEN technical ecosystem was preceded by a model of the social media environment that could be used by software designers to ingest social media information and make predictions about social media users and their targeted audiences. This model separated the social media environment into the narrative (i.e., message content) component and the community (i.e., individuals, groups, networks) component. In a white paper prepared at the Office of Naval Research (ONR), its researchers proposed that by modeling these two components, an analyst or planner with appropriate tools and methods could ascertain the intentions of social media campaigns designed to influence behavior (Carley, 2022). This near-real time monitoring can be added to an IE analyst's understanding of the overall situation, including the language, culture, and history of the targeted audience, providing a relevant assessment of the impact of social media on behavior. Using purpose-built analysis software developed to ascertain both the content of the messaging and a community of users provides a means of quickly monitoring social media and its potential effects on military operations and broader population.

Information Environment

The IE is a consideration in all military operations with Operations in the IE (OIE) being described as imperative at every level of war in Joint and Service doctrine (Joint Chiefs of Staff, 2022a). Over the past thirty years, the IE has become an increasingly larger component of all operations. The IE and its physical, information, and cognitive (human) components have been impacted by the tremendous growth in volume of information available delivered at near-real time speed (Joint Chiefs of Staff, 2022a). Information networks have grown to reaching almost anywhere in the world, in part because of the ubiquity of handheld devices. The 'will' of the populace has always been a factor at the strategic level of war, but with this real-time information

dissemination at the individual level, every person affected by military operations possesses the potential to affect outcomes. This path of information can be exploited by friendly and malign actors alike to achieve information advantage on networks or channels chosen by the disseminator. The military has leveraged information and their understanding of its effects over the centuries to create deceptions or influence an opposing military to take (or not) action to achieve information advantage and alter an opponent's chances of success. Similarly, US forces have a long history in the development of techniques to influence behavior, attitudes, or perceptions of both combatants and non-combatants in an area of operations. The volume of information and the speed with which it is disseminated have created a requirement for commanders at every level to consider the role of social media as both a threat and opportunity to influence behavior of those affected by military operations. Therefore, the IE must be described just as every other domain is described in military planning. Threats to the success of military forces must be determined and plans proposed that minimize or counter such threats.

One component of the IE is the various social media platforms. Information collected from social media largely falls into the category of Publicly Available Information (PAI) and/or Open-Source Intelligence (OSINT) and represents a daunting environment for even experienced analysts that often must consume hundreds of reports per day. Social media content generated by individuals or intelligent agents like bots and cyborgs deployed within the area affected by military operations could easily overwhelm tactical or operational level intelligence analysts who lack the automated means to process them. Further, social media messages that emanate from outside the area directly affected by military operations but are designed to influence behavior can number in the millions. Only purpose-built analysis software can enable an analyst to quickly 1) identify the community (e.g., networks) of those that are attempting to influence behavior along with the individuals and groups targeted and 2) compare the content (e.g., topics, sentiment) of each social media message to assist the analyst in identifying the purpose of each issuance.

Information that can be collected through open sources, such as online social media, printed media, broadcast media, individual smart phones, is virtually limitless and can be analyzed to inform influence operations. The dissemination of information, especially when it reflects local context and culture, can have an immediate effect on behavior in an area affected by the conduct of military operations. Rapid information dissemination creates opportunities for malign actors to potentially affect behaviors, attitudes, or perception in an area that would impact US and allied operations before friendly forces react. Therefore, friendly forces are required to leverage cutting edge technology to maintain or restore their information advantage.

The connection between online behavior and human actions in the 'real world' may often be challenging to detect but this is exactly what the military analyst is trying to find. Within the information flow are cues and indicators of group identity and networks that suggest how messages are used to target populations, including changes or roles within networks, sentiment, or social network engagement patterns. In effect, messaging is part of operations that support the command narrative. The use of social media is an added factor that must be integrated with the decision calculus designed to achieve an effect in the IE.

In the multi-domain battlefield, the competition is often between rival networks whose size and composition might be difficult to visualize. Data-driven solutions and advanced technology are necessary but insufficient for achieving effective performance. Human assessments supported by advanced technology and software can be used to provide insights about effectiveness. Assessing the effectiveness of information operations in the internet mediated communications world presents a complex problem. The challenges of providing a reasonable estimate of effectiveness include the dynamics of the IE, the vast volume of information that must be curated and managed, the increasing velocity of information flow into and through the system, and rapidly evolving nature of trends and patterns including changing social media platform algorithms. In this project, we focused on how human performance contributes to achieve information dominance by defining measures, digitizing the collection process, and rapidly synthesizing information and confirming judgments to evaluate individual and team performance. Because measuring individual performance and team effectiveness is a complex problem, the research team sought an empirical, reliable, straightforward approach to measurement. The approach used for the current effort was technology agnostic because systems and technologies are ever-changing as society advances; we assumed that the users would adapt to these innovations when necessary and required.

Behaviorally Anchored Rating Scales

The foundation of BARS for assessment had the goal of enhancing and standardizing observations rather than using a retrospective approach to assessment, which often leads to bias and error (Smith & Kendall, 1963). BARS are designed to be objective and context-specific to assess behaviors along specified dimensions (Bernardin & Smith,1981). BARS move away from judging performance based on traits to a construct that captures the behaviors and processes demonstrated by individual performers (Schwab et al.,1975). Using BARS, a standardized set of measures, supports reliable data collection and meaningful

performance assessment focused on cognitive performance and mental models rather than psychomotor skills or declarative knowledge. BARS have been successfully applied in a multitude of domains, including nursing (Smith & Kendall, 1963) and instructional leadership and facilitation (Phillips & Ross, 2016).

One commonly used structure for BARS is each item is an individual task and that task includes a five-point nominal scale consisting of performance indicators. This five-point nominal scale can be based on the five stages of development ranging from novice to expert that was proposed by Dreyfus and Dreyfus (1980; 1986; Figure 1). This five-level structure has been applied in cognitively complex domains where decisions are made in dynamically complex environments that involve risk, uncertainty, and ambiguity (Ross & Phillips, 2020). The characteristics of performance distinguishing the five stages of development have proven consistent across domains. When structured this way, the BARS includes higher-order categories for the individual tasks called key performance areas (KPAs). Each KPA will have a series of associated tasks or broad behaviors. Each task will have five associated behavioral indicators that represent increase in performance from novice to advanced beginner to competent to proficient to expert).

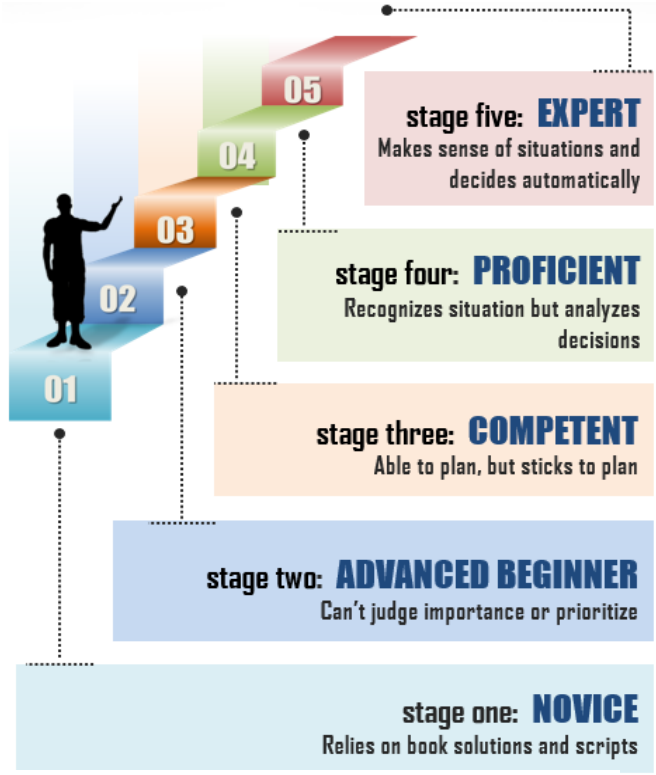

Figure 1. Five Stages of Development

Field Assessment System

Using the Field Assessment System (FAS) for rating performance using the BARS items is a more agile approach for data collection than relying on a paper-and-pencil version. After setting up the FAS with the relevant BARS rubric, the FAS will organize the screen to present all KPAs and the MOPs that fall within each KPA. This makes it easier for raters to focus on only relevant MOPs by searching within the appropriate KPA. Trained raters can view and score the MOPs on the main screen, depending on the number of MOPs and KPAs, rather than browsing through multiple pages to find the necessary MOP. Simply selecting the appropriate MOP with a stylus or finger opens the dialogue box where the rater selects the corresponding behavior/rating (Figure 2). Additionally, the FAS allows raters to add comments on the overall rating session and to make comments on specific ratings to provide additional detail as to why the given rating was selected. This keeps notes organized and situated within the appropriate context.

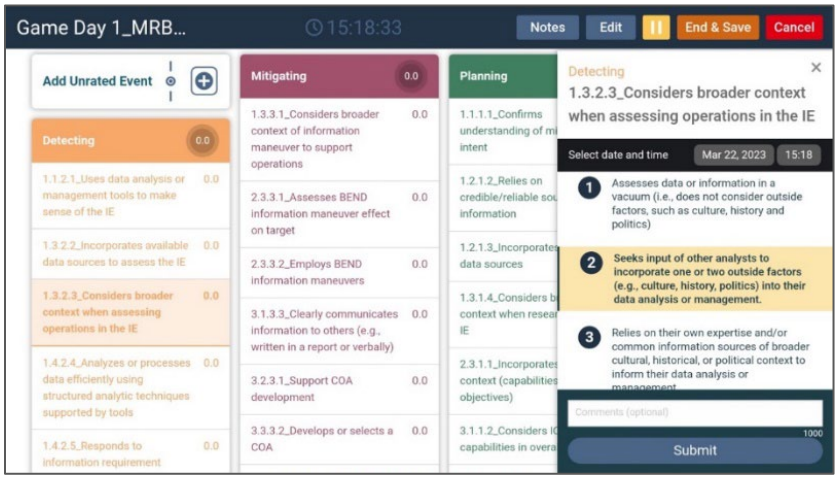

Figure 2. The Rating Dialogue Box on the FAS

The FAS is an application that operates on a tablet and is usable in a laboratory or field setting. The FAS consists of a user interface, a data structure, item definitions, and a scoring rubric (i.e., BARS items and associated behaviors and scores). The FAS stores performance data collected by each rater for the rating session (e.g., a full exercise or a subset of an exercise) and maintains a time data of for each rating that allows the rater to re-trace the event and provide post-event comments. All the data collected and stored on the FAS and tablet can be exported as a .CSV file, which maintains the KPA-based structures implemented in the pre-defined rubric. Data can then be moved to different software for deeper analysis, such as Power BI.

CURRENT EFFORT

The current work aimed to address the measurement issue by developing a measurement method that is widely applicable, technology agnostic, and user-friendly.

Method

The method used to develop the O-BARS included simplifying the tasks and associated behavioral indicators, making the MOPs more user-friendly, and digitizing the BARS. This method was straightforward and general so that it could easily be applied to other domains resulting in new BARS for those domains or jobs. Each of these aspects of developing the current version of the O-BARS will be described in turn.

Developing a User-Friendly O-BARS

To create the first version of the O-BARS, we started with the Information Maneuver Analyst (IMA) Mastery Model (Borders et. al 2022). The IMA Mastery Model was developed previously using five data sources: (1) the Intelligence, Surveillance, and Reconnaissance (ISR) Mastery Model (Phillips et al., 2019), a framework for assessing career development, training assessments, and performance critiques of the Intelligence community; (2) canonical student materials, training and rehearsal curriculum for introducing social media analysis technologies and guided practice and feedback using a social media database; (3) the after-action review notes from a previous OMEN related military exercise; (4) the BEND Playbook, a primer to guide planning, execution, and assessment of social media behaviors and maneuvers (Goolsby & Carley, 2019); and (5) informal discussion with practitioners using the relevant behaviors, tools, and training. The IMA Mastery Model included five KPAs with a total of 106 items. Transitioning each mastery model item into an MOP for the BARS would result in an O-BARS with 106 tasks, which is unwieldy and unusable. Rather than a one-to-one transition from the IMA Mastery Model to the O-BARS, we combined and parsed down the data into broader, collective items resulting in a user-friendly solution that was easily implementable in an exercise environment or demonstration.

The first step of this process was to sort each task from each KPA into three stages of the pattern of operations previously identified: planning, operations, and assessment. Planning is the continuous, cyclical process of making decisions and conducting operations for a specific objective by defining the problem, understanding the adversary, the environment, the friendly situation, and mission requirement (Joint Chiefs of Staff, 2020). Operations is the adaptive process of applying what is known about the IE and mission objectives to the implementation of a course of action (COA) for achieving the desired end-state within constraints based on an understanding of own capabilities and an adversary's intentions (Joint Chiefs of Staff, 2022b). Assessment is the continuous, systematic process of surveilling and making sense of the IE by gathering and processing data, perceiving patterns or anomalies of interest, and generating data-informed assessments for predicting a future state based on a reliable method of analysis in support of the mission (Joint Chiefs of Staff, 2022a).

The second step of creating a succinct O-BARS, we used a card sorting method where all tasks in each KPA and process step were sorted into larger themes with multiple tasks falling into each category. Each broader category was renamed with a summarizing task; this became the new MOP. Each BARS MOP targets individual thinking skills and mental models that an information maneuver analyst applied to ill-structured domains like social media.

The third step, once all tasks within each KPA had been sorted and summarized, involved rewriting each behavioral indicator for all five levels (i.e., novice, advanced beginner, competent, proficient, expert) in every MOP. This ensured that all the original behavioral indicators within each task from the mastery model were represented in the new behavioral indicators within each MOP. In a second pass through the behavioral indicators, we modified the language to ensure consistency, coherency, clarity, and conciseness, and to ensure that all behaviors were truly observable and did not require making inferences about intent. Although the O-BARS were finished at this stage, we evaluated the relationships between the subcategory level of the O-BARS and Training and Readiness manuals (Department of the Navy, 2011, 2022) to add additional rigor and validity.

This version of the O-BARS resulted in 59 total MOPs across the five KPAs. Figure 3 shows a screenshot of a paper-based version of the O-BARS. Figure 4 shows the number of MOPs for subcategory within each KPA.

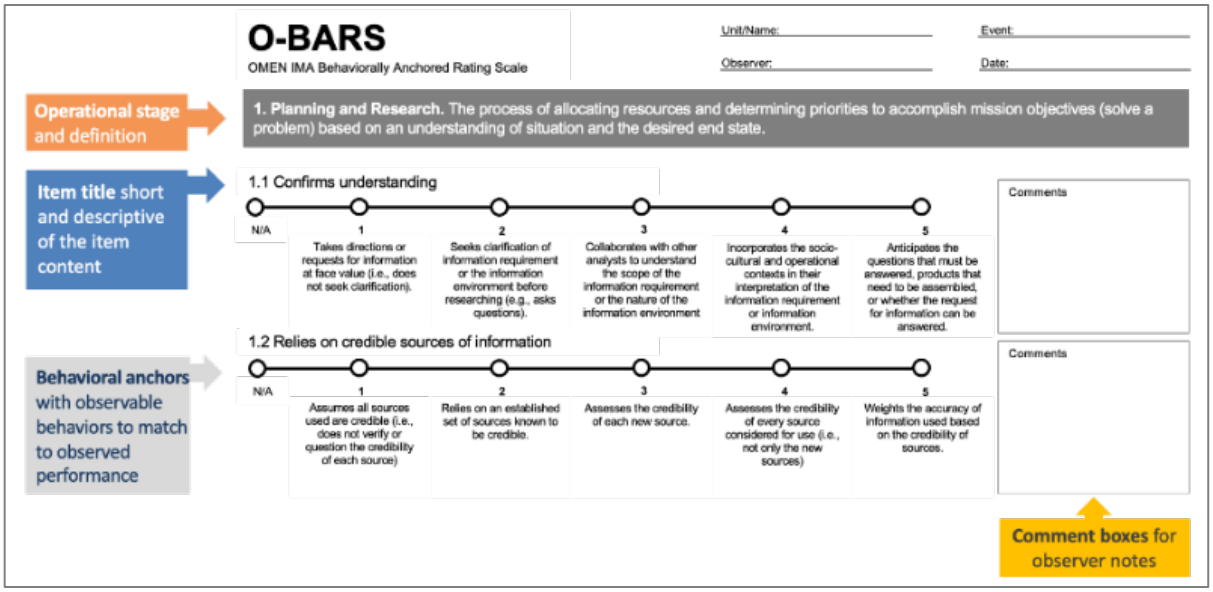

Figure 3. Paper-Based Version of O-BARS

KPA 1.0 Information Gathering (10)		KPA 2.0 Analysis & Synthesis (14)		KPA 3.0 Integrating Intelligence (13)		KPA 4.0 Maintaining Awareness (7)		KPA 5.0 Reasoning & Seasoning (15)	
Subcategory	MOPs	Subcategory	MOPs	Subcategory	MOPs	Subcategory	MOPs	Subcategory	MOPs
1.1 Understanding the Information Requirement	(2)	2.1 Determining the Analytic Method & Technique	(4)	3.1 Performing Operators' Requirements	(6)	4.1 Maintaining Battlespace Awareness	(2)	5.1 Attending to Cognitive Bias	(3)
1.2 Identifying Data Sources	(2)	2.2 Sensemaking and Drawing Conclusions	(5)	3.2 Applying Intelligence to Information Operations	(4)	4.2 Updating Knowledge of the Adversary	(2)	5.2 Problem Framing	(3)
1.3 Judging Information Quality & Credibility	(4)	2.3 Identifying Threats & Adversary COAs	(5)	3.3 Supporting Mission Planning	(4)	4.3 Processing Incoming Information	(3)	5.3 Questioning Facts & Assumptions	(5)
1.4 Assessing Data Relevance	(2)	2.4 Leveraging Automation	(1)					5.4 Judging Results	(7)

Figure 4. Number of MOPs Per KPA and Subcategory

Constructing Measures of Effectiveness

Measures of effectiveness (MOEs) are a measure of a system's capability to achieve an objective in a particular condition. For the purposes of this effort, we took the perspective that the IO cell is the system, and the IE is the condition. Therefore, MOEs were focused on whether the analysts and planners were performing the tasks correctly, so that whether the IO cell performance could be assessed. Relying solely on individual performance measures (i.e., MOPs) to inform the assessment of IO cell performance might result in gaps, subjective judgment, and logical fallacies. As a first step towards an MOE solution set, we developed a doctrinally grounded view of team effectiveness (Department of the Navy, 2011, 2022; United States Marine Corps, 2022).

After a set of MOEs were defined based on doctrine, the pre-developed MOPs were sorted into the MOEs using a card sorting technique. The process for determining the MOEs involved three steps: 1) three researchers separately sorted the MOPs into the MOEs representing a many to many relationship, 2) the three researchers met on multiple occasions to discuss disagreements on categorization relying on doctrine, and 3) the researchers came to a consensus based on discussions to create a final version of MOEs and associated FMOPs. This resulted in eight doctrinally-based MOEs with face validity that can be tested in an operational setting.

Digitizing BARS Using FAS

With the large number of MOPs, digitizing the O-BARS served to make the measures more usable and agile. Using a digitized system instead of a paper and pencil approach would result in efficient data collection, focus collection on participant performance, and simplify post-processing and analysis. The O-BARS was digitized using a tool

previously developed under ONR—the FAS. The FAS satisfied the requirement for a flexible tool suitable for collecting, compiling, and scoring individual performance.

The FAS has been used to successfully digitize BARS beyond the O-BARS. For our purposes, we modified the BARS structure to fit the requirements of the FAS so that the O-BARS rubric could be uploaded rather than manually created within the FAS. After data are collected for that person and event, the FAS provides a quick view of the results of the selected rating session on the results tab (Figure 5). The FAS results tab is formatted to set the X-axis with the next highest whole number rather than the highest possible rating. For example, if the highest rating is 2.7, the x-axis maximum is 3. Researchers can export the data from FAS for more in-depth analysis and data visualization using Power BI (see Figure 6 for example visualizations), which allows for customization ranging from colors of graphs to x and y-axis maximum values. Using Power BI to develop clear and accurate data visualization supports data-informed decision making.

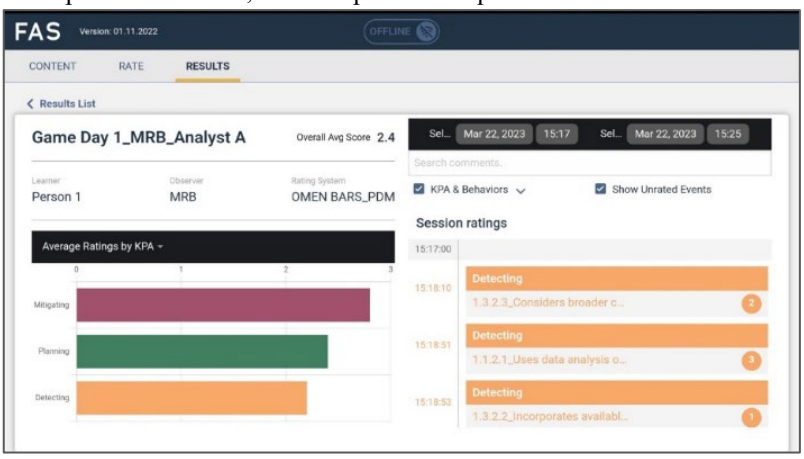

Figure 5. Example Results Screen on the FAS

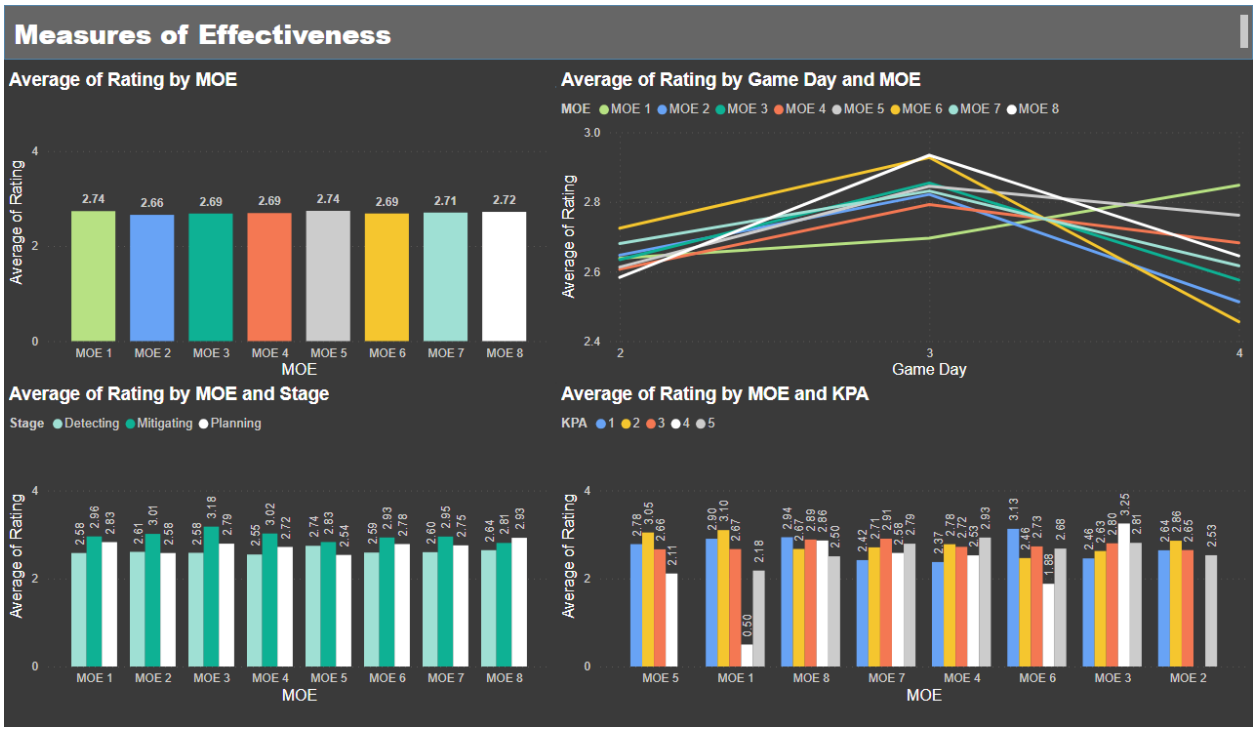

Figure 6. Example Power BI Visualization of O-BARS Data

Note. The Y-axes range from 0 to 5 because MOP ratings can be scored as 0 (not rated) through five. However, the top right figure ranges from 2.4 to 3.0 so that the lines are easily distinguishable for the purposes of this paper. This can be manually changed within Power BI to reflect the true rating scale and to accurately represent the data. Further, most ratings depicted in all the graphs range from approximately 2.00 to 3.25 indicating most performers were rated as advanced beginners with a few rated as competent. This suggests that performers were likely following the provided rules and procedures without deviating.

Proof of Concept Demonstration

The O-BARS rubrics were tested at an OMEN Demonstration at Carnegie Melon University. This technical demonstration supported the OMEN program's objective to demonstrate progress towards a fully integrated OMEN solution for social media analysis and performance feedback. A scenario-based technical demonstration of social media analysis tools was conducted to make sense of the IE in the lab.

During this event, we demonstrated and tested the use of the O-BARS for assessing IO cell performance, specifically analysts and planners. More specifically, we verified the suitability of the MOPs (e.g., the content, the design, the descriptions) and the usability of the digitized version.

Twelve participants worked through a multi-day OMEN-team generated scenario using the various OMEN tools and technologies. Participants were divided into two groups; each group was working on the same scenario but separately. The event took place across 3 days with a total of four "game days."

Each game day followed a similar pattern of operations. Following the orientation period, two teams were assigned roles as analysts and planners (i.e., there was no explicit distinction between the two roles; all performers acted as both analysts and planners). The White Cell communicated task orders or events for execution. Each task order provided a mission statement and objectives. Participants were oriented on task order objectives issued by the White Cell. The pattern of operations is illustrated below (Figure 7) with three stages (i.e., Planning, Operations, and Assessment). The effects were adjudicated and set the conditions for the subsequent game day.

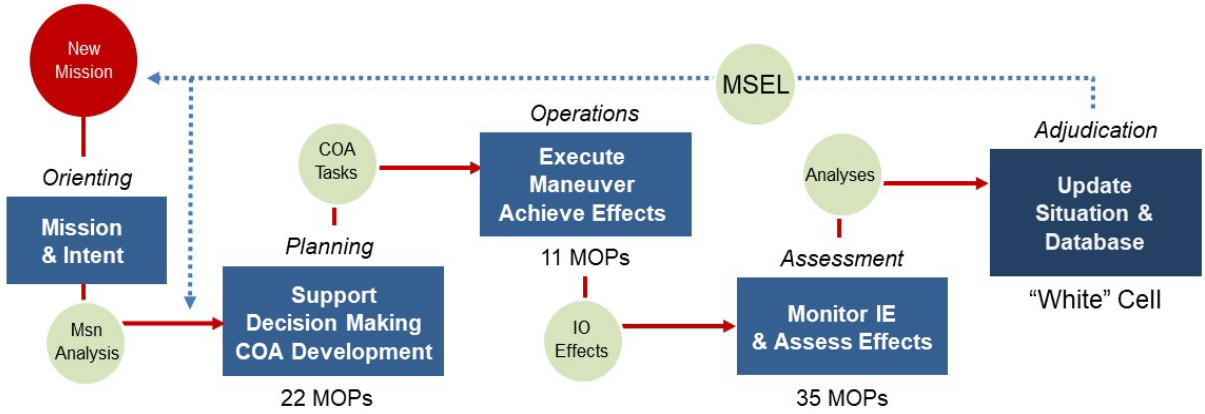

Figure 7. OMEN Demonstration Pattern of Operations

Note. MSEL stands for Master Scenario Events List; MSN stands for mission.

Three raters observed participants as they worked through the scenarios to rate them using the O-BARS on the FAS. Raters observed a different participant for each game day ensuring that the O-BARS applied across individuals regardless of experience and background, and therefore was more generalizable. Additionally, each game day raters switched from which group they observed a participant. Game Day 1 served as a calibration day for the O-BARS raters to ensure each team member understood the exercise, how to use the FAS, and the nature of the simulated performance environment, including the pattern of operations. Raters collected usable data on Game Days 2-4.

Lessons Learned

Testing the O-BARS during the technical demonstration proved successful and beneficial. However, because the purpose of the test was to determine whether the O-BARS and the FAS are appropriate for the assessment of analyst and planner performance, this section discusses the lessons learned rather than the ratings of the performers. We identified general findings and specific takeaways that will inform future iterations of the O-BARS.

General Findings

During the technical demonstration exercise, the O-BARS on the FAS was successfully implemented by all raters regardless of technological literacy or experience with the FAS. Raters were able to initiate and use the FAS, navigate the BARS, and rate participant performance without difficulty after the calibration period. The MOPs used during this

demonstration were appropriate and raters were able to identify and rate the described behaviors. Further, using Power BI for additional analyses and visualizations was a user-friendly way to share the findings with others and would clearly serve as an effective source of information for data-informed decision making. Overall, the researchers were able to achieve their goals for the demonstration and test the O-BARS in a simulated environment.

A major benefit of testing the O-BARS in this environment is that we were able to calibrate to ensure all raters were prepared and agreed on the method. Additionally, we were able to interject and ask participants to walk us through what they were doing so that we gained a better understanding of their task and the technologies. Further, the research team was able to use this opportunity to inform future iterations of the O-BARS.

Specific Takeaways and Actions

As part of the research objectives, we aimed to identify areas of improvement in terms of content, structure, process, and technology. Our current areas of improvement are as follows:

1. The efficacy of individual analyst and planner performance measurement depended on the clarity of the item as well as the ability of the observer to recognize the behavior so it could be rated using the behavioral indicators. For O-BARS to be effectively implemented, we learned it will be necessary to revise the item descriptions as the next step in adopting this approach.
2. Understanding the context of IO was essential to providing meaningful performance ratings. Preparing data collectors and researchers to apply O-BARS effectively requires a knowledge of information maneuver doctrine as well as hands-on experience with the FAS.
3. Data structure based on key performance areas facilitated data collection and streamlined the post-processing and analysis activities. Applying the 3-stage framework and sequencing the performance measures within each stage allowed the observers to focus on the performance of interest more consistently.
4. The initial O-BARS was demonstrated in a limited objective demonstration to assess how analysts and planners made sense of the IE to apply social media platforms to influence a target audience's behaviors, attitudes, and perceptions. Revising O-BARS based on the initial results indicates that the rubrics are easily adapted to the full range of information related capabilities.
5. Demonstration of the O-BARS during a limited objective technical demonstration confirmed that the FAS is a more efficient collection method compared to pen and paper. However, the results did not verify the suitability of the items. Verification of the O-BARS measures during information warfare exercises has been coordinated and should occur in calendar year 2024. Collaboration with the operational community to verify and validate the measures of effectiveness for assessing IO cell collective performance is also essential.

Since the demonstration, the research team made several of improvements to the current version of the O-BARS. The MOP descriptions have been updated so that they are more descriptive and discriminant. The existing training on the FAS and O-BARS has been updated and will continue to be refined with testing. The researchers have developed a new rubric that has been uploaded to the FAS with the stage as the highest-level organizational factor rather than KPA. At this point, it appears as though this organizational scheme will prove beneficial for implementing the O-BARS in the field. Eight new MOPs have been added to the existing MOPs and the data structure has been updated to include the KPA, subcategory, stage, and MOP number. Further, additional means to streamline the data analysis process are underway with the development of a system to automatically reformat the FAS data output for Power BI analysis and additional Power BI templates. Future work will aim to test and verify the O-BARS, including the MOEs and MOPs, in the field and with the relevant groups.

BENEFITS AND PAYOFFS

O-BARS offers a scientifically grounded, data-informed approach to assessing individual and collective IO planners' and analysts' performance that is prescribed in the Information Maneuver Training and Readiness manuals (e.g., Department of the Navy, 2022). O-BARS complements current solutions for preparing teams of planners, analysts, and decision makers, who must make sense of data and patterns gathered from online platforms. The results point to the need for training solutions that apply data science and critical thinking skills. O-BARS provides the yardstick for objectively gauging what is working well and where the Services should invest in training and exercise solutions.

Because of the objective nature of the O-BARS, the behavioral indicators provide a means for avoiding bias and limiting subjectivity when evaluating individual or team performance.

Although we targeted analysts and planners who deal with information maneuvers and social media data in the current application, the O-BARS can play a broader role for evaluating the cognitive performance of analysts and planners in other information related capabilities. Because other IO planners and analysts combine human-cognitive performance with advanced data management tools and technologies, an updated version of the O-BARS could apply. Some adaptation would be necessary to provide reliable performance assessments for other elements of a Marine Expeditionary Force (MEF) Information Group including Military Information Support Operations (MISO), Communications Strategy and Operations (COMMSTRAT), Public Affairs, Intelligence, Cyber and Fire and Effects Coordination. Any modifications would likely be minor, and the transition would be simple.

The O-BARS ratings can be applied to discern performance trends and insights about professional development needs. O-BARS data represent a virtual snapshot of individual performance, team effectiveness, and the sources of friction or error in performance. This approach recognizes the nature of performance is interdisciplinary and steeped in detail about critical thinking, error detection and correction, achieving situational understanding, and sensemaking. The O-BARS yields the human-system requirements necessary for improving and directing team performance by focusing on system operators instead of their tools.

Further, in addition to using the O-BARS to identify professional development needs, training developers can leverage O-BARS data to target gaps in human performance (individual or team) by highlighting areas for improvement. Because the performance ratings are available quickly, they support formative feedback and coaching during a training session. This adult learning approach of quick and targeted feedback supports transfer of learning to enhance performance that is observable and repeatable on the job. Training developers could enhance or modify training plans based on O-BARS data to address consistent performance gaps across multiple organizations or teams. Similarly, because IO effects might not be evident or provide convincing results in the short term, O-BARS data could contribute to the adjudication of information effects on operations. O-BARS used to shape scenario events would be the results of compiling results from previous exercises or experiments and using the information for exercise and training design. Therefore, the O-BARS can be used to detect larger systematic gaps in training effectiveness and be used to inform scenario events.

REFERENCES

Bernardin, H.J., & Smith, P.C. (1981). A clarification of some issues regarding the development and use of behaviorally anchored ratings scales (BARS). *Journal of Applied Psychology, 66*(4), 458–463. https://doi.org/10.1037/0021-9010.66.4.458

Borders, M.R., Ross, W.A., Williams, M.L., McGee, R., & Danks, D (2022). OMEN: Final Technical Report. (Technical Report prepared under contract N00014-21-C-2045 for the Office of Naval Research). Portsmouth, VA: Cognitive Performance Group.

Carley, K.M. (2022). The Digital Social Media Playbook 3.0 [Unpublished manuscript].

Department of the Navy. (2011, April 5). *Public affairs (PA) training and readiness (T&R) manual.* (NAVMC 3500.9A). https://www.marines.mil/Portals/1/Publications/NAVMC%203500.9A.pdf

Department of the Navy. (2022, October 18). *Information maneuver training and readiness manual* (NAVMC 3500.124A Ch 1). https://www.marines.mil/Portals/1/Publications/NAVMC%203500.124A%20Ch%201.pdf?ver=QfRLAaHhO4nqcyjdo_MamA%3d%3d

Dreyfus, S.E., & Dreyfus, H.L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. California Univ Berkeley Operations Research Center.

Dreyfus, S.E., & Dreyfus, H.L. (1986). *Mind over Machine.* New York, NY: Free Press.

Goolsby, R. & Carley, K. (2019). *The Digital and Social Media Playbook Concept of Operations (v1.5).* FOUO. White Paper, Office of Naval Research

Joint Chiefs of Staff. (2022a). *Information in Joint Operations* (JP 3-04).

Joint Chiefs of Staff. (2022b). *Joint Operations* (JP 3-0).

Joint Chiefs of Staff. (2020). *Joint Planning* (JP 5-0).

Phillips, J.K., Hancock, A.K., Baxter, H.C., Williams, M.L., Pfahler, M.D. (2019). U.S. Marine Corps Intelligence, Surveillance, and Reconnaissance Mastery Model. Developed for Office of Naval Research, Arlington, VA under Contract N00014-18-C-2034.

Phillips, J.K., & Ross, K.G. (2016). Option III: Instructor assessment battery development report. (Technical Report prepared under Contract N00014-14-C-0106 for the Office of Naval Research). Orlando, FL: Cognitive Performance Group.

Ross, K.G., & Phillips, J.K. (2020). Developing mastery models to support the acquisition and assessment of expertise. In P. Ward, J.M. Schraagen, J. Gore, and E. Roth (Eds.), *Oxford Handbook of Expertise* (pp. 312-332). Oxford: Oxford University Press.

Schwab, D.P., Heneman, H.G., & DeCotiis, T.A. (1975). Behaviorally anchored rating scales: A review of the literature. Personnel Psychology, 28(4), 549–562. https://doi.org/10.1111/j.1744-6570.1975.tb01392.x

Smith, P.C., & Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 47(2), 149–155. https://doi.org/10.1037/h0047060.

United States Marine Corps. (2022). *Information* (MCDP-8). https://www.marines.mil/Portals/1/Publications/MCDP%208.pdf?ver=6gIvEcD0CUuPAgTSmyDNag%3d%3d