Development of Key Variance Visualizations of Analytic Workflow for the Support of Data-Based Discussions

Nicole M. Francavilla (nfranca@ncsu.edu), Samuel Wilgus (sjwilgus@ncsu.edu),
and Mark A. Wilson (mark@ncsu.edu)
NC State University Laboratory for Analytic Sciences

The Problem

Analysts have a diverse set of responsibilities, making it difficult for branch leaders to identify behavioral areas to target when tasked with improving branch performance. While there are systems in place to evaluate performance and provide analysts feedback, these systems may lack focus or efficiency and fail to provide guidance on how branch leaders should use the information they provide. The intelligence community would benefit from a system that directs branch leaders' attention to performance areas worth addressing.

If you were to survey a group of five employees that complete the same task daily, they might provide five different ways to complete the task. We would call this diversity *variance in workflow*. The presence of variance itself, however, does not necessarily warrant concern. Another piece of information worth evaluating is the extent to which the variance affects analyst performance. It is those behaviors that have high variability and also affect performance that deserve a branch leader's attention, and the intelligence community does not currently have a system in place that can quickly identify such areas.

Failure to optimize the discussion within teams, branches, or divisions can result in significantly lower performance, cohesion, satisfaction, innovation, and knowledge integration (Mesmer-Magnus & DeChurch, 2009; Wang & Wang, 2012). Within the intelligence community, low levels of cohesion, reduced performance, and limited knowledge integration could prevent talent leverage and lead to critical mistakes or oversights. Optimizing discussion and information-sharing could benefit the intelligence community by facilitating development, innovation, integration, and performance within work groups.

Unfortunately, neglecting discussion structure within groups is a major impediment in optimizing interaction and collaboration. Neglecting discussion structure can range from failing to pool information (Wittenbaum & Bowman, 2004) or lacking structured discussion procedures (Strasser, Taylor, & Hanna, 1989). Lack of discussion structure translates into a lack of shared knowledge and information, which can negatively influence the group's problem solving ability

(Lu, Yuan, & McLeod, 2012). Therefore, any attempt to optimize communication and discussion within a group must also aim to improve the structure of the group discussion.

In regards to analytic workflow behaviors, a lack of discussion structure and information-sharing may result in misconceptions about the ideal frequency for various important behaviors. For example, individuals might assume it is always best to simply capitalize on strengths and avoid weaknesses, especially during periods that require maximum performance. Leadership research has found that doing too much of a certain behavior (even a behavior that is typically seen as positive, such as being sensitive) can be just as detrimental to performance as doing too little (Kaiser & Hogan, 2011). A critical step to achieving this type of balance within analytical workflow behaviors is through establishing a system of discussion structured on behaviors that exhibit key variance (doing too much or too little). The present study implemented a performance evaluation system that measured the extent to which analytic workflow behaviors were both underdone and overdone, and also measured whether they affected performance. The ultimate goal for the use of this data was to guide data-based discussions between branch leaders and their analysts.

The Field Study

Twenty-four analysts from the same branch of one intelligence agency participated in the field study. Due to the sensitive nature of these data, exact return rates will not be reported; however, we are confident that this sample is representative of the branch as a whole. These data were collected during surge operating conditions and participants represented a range of organizational and in-role tenure. About half the sample reported being a subject matter expert and 40% reported being a leader.

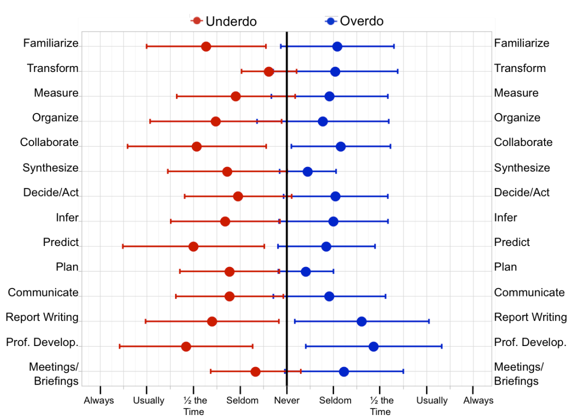
The field study consisted of a key variance survey administered by a trusted analyst in one-on-one intake sessions. The survey contained two intake forms, one for normal operating conditions and one for surge conditions. On both forms, participants provided ratings on fourteen dimensions of analytic workflow behavior. Example dimensions are “Familiarize” and “Predict.” This set of dimensions was derived via observations of, and interactions with, intelligence analysts. Dimension development was an iterative process and experienced analysts’ reactions to the chosen set were positive.

The key variance forms collected two types of information about each behavioral dimension: frequency ratings and performance ratings. For frequency, participants rated the extent to which their peers underdo and overdo each behavior. For performance, participants rated whether under- and overdoing each dimension ultimately affects performance. In sum, on each of the two key variance forms (normal and surge), for each dimension, participants provided four ratings: an underdo frequency rating, an underdo performance rating, an overdo frequency rating, and an overdo performance rating.

The Results

The key variance analysis process aimed to address five research questions, outlined below. We analyzed the normal and surge data separately. While participants provided data on their peers, they reported that the set of dimensions captured 93% of their personal activities.

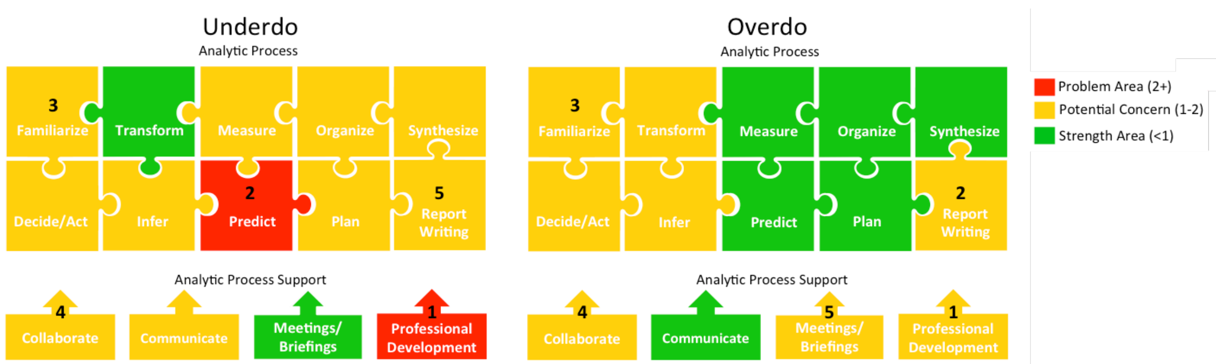
Question 1 asked: *Is there variance in analytic workflow?* For both the normal and surge data, we computed means and standard deviations for each dimensions' underdo and overdo frequency ratings and plotted this information to visually examine the extent of variance. The graph for normal conditions is to the left. Ratings surrounding the zero line represent those dimensions not perceived to be under- or overdone frequently, and ratings that move further away from the zero line represent those perceived to be under- or overdone more frequently. The graph indicates that variance certainly exists in analytic workflow, as means vary across dimensions. For instance, some behaviors were perceived to be seldom underdone, while some were perceived to be underdone about half the time. Additionally, the long error bars represent variability in ratings, meaning that participants had some disagreement on each dimension's frequency.



Question 2 asked: *Does workflow variance affect performance?* We calculated the percent of participants who indicated that under- or overdoing each dimension affects performance. Percentages ranged from 20% to 77%, indicating that some dimensions were perceived to affect performance more than others across the sample.

Question 3 asked: *Where is the key variance?* We synthesized the information provided by the frequency and performance ratings in order to identify the areas of key variance in analytic workflow. To do so, we rank ordered the dimensions by their frequency mean and performance percentage separately. Then we took the average of these two rank scores for each dimension, and re-ranked them based on this value. This technique identifies those behaviors that are both under- or overdone frequently *and* impact performance. We compared the top five areas of key variance for normal and surge conditions, and found more similarities than differences, with the two contexts sharing four out of five key variance areas for underdo, and three out of five for overdo, indicating that some behaviors are perceived to be under- or overdone often, and affect performance, regardless of the work context.

In order to visually summarize this information in a straightforward format, we developed key variance visualizations for managers (see below). These visualizations are color coded based on mean frequency ratings and numbered based on the top five areas of key variance. Such reports signal to managers which behaviors need attention. The numbered dimensions are those which we suggest data-based discussions surround, and the color coding



helps to prioritize which are most problematic. Based on the visualization below, the manager should highly prioritize discussions about underdoing Professional Development and Predict.

Question 4 asked: *Are there meaningful subgroups of analysts?* We conducted two separate cluster analyses on the normal and surge frequency data. This approach places participants into subgroups of individuals who had similar patterns of responses. We identified 4-cluster solutions for both conditions, meaning there were four unique subgroups of analysts based on under- and overdo behavior. However, many participants did not fall into the same subgroups for normal and surge conditions.

Question 5 asked: *How do the subgroups differ?* We compared these empirically derived groups to rational groups created based on work roles and found many dissimilarities, indicating that interventions targeting work groups may not be effective. Additionally, we compared the subgroups to one another within the normal and surge conditions, and found that they represented different types of performers; for instance, one group's mean frequency scores hovered around the zero line, indicating that they under- and overdo behaviors infrequently, while another group consistently had high ratings for underdoing behaviors. These group differences provide evidence that a one-size-fits-all key variance approach may not be most effective, and interventions may be better targeted to smaller pockets of similar performers.

The Road Ahead

Subject matter expert feedback on this approach indicates that it may play an important role in facilitating data based discussions to improve analytic workflow. Results reported here need to be replicated in other branches and extended to roles other than analysts. More work is also needed to better understand how to conduct data based discussions of workflow.

We hope to collect more data in other branches to determine whether the results replicate or if key variance areas differ across branches. Similarities would indicate that key variance results more from analytic workflow in itself, whereas differences across branches would indicate that key variance results from contextual variables such as leadership styles, training opportunities, and other aspects of the work environment. In addition to collecting more data on analytic workflow, we would like to extend key variance investigation to managerial, researcher, and developer roles. This would entail development of behavioral dimensions that are representative of the role of interest followed by data collection in branches of these workers. One challenge we face with data collection is the importance of anonymity to participants. When anonymity is issued, it follows that data-based discussions can only address performance of the group as a whole rather than identified subgroups or individuals.

We also aim to develop more specific guidelines on how to conduct data-based discussions that leverage the information provided by key variance visualizations. We intend to expand on what these discussions might look like (i.e., workshops, weekly meetings, online communications, etc.) and how managers can effectively work with their employees to identify both determinants of key variance areas and solutions to address them. In the future, our goal is to have an automated system that can input a branch's key variance ratings at a given time and provide a key variance visualization report which compares back to past intakes, monitors progress for that branch, and normatively compares it to others.

References

Kaiser, R. B., & Hogan, J. (2011). Personality, leader behavior, and overdoing it. *Consulting Psychology Journal: Practice and Research*, *63*(4), 219.

Lu, L., Yuan, Y. C., & McLeod, P. L. (2012). Twenty-five years of hidden profiles in group decision making a meta-analysis. *Personality and Social Psychology Review*, *16*(1), 54-75.

Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: a meta-analysis. *Journal of Applied Psychology*, *94*(2), 535.

Stasser, G., Taylor, L. A., & Hanna, C. (1989). Information sampling in structured and unstructured discussions of three-and six-person groups. *Journal of Personality and Social Psychology*, *57*(1), 67.

Wang, Z., & Wang, N. (2012). Knowledge sharing, innovation and firm performance. *Expert systems with applications*, *39*(10), 8899-8908.

Wittenbaum, G. M., & Bowman, J. M. (2004). A social validation explanation for mutual enhancement. *Journal of Experimental Social Psychology*, *40*(2), 169-184.