EasyChair Preprint

№ 5830

Responses to Overload as Foundation for a Resilient Project and Organizational Health Metric

Marisa Bigelow, Lisa Douglas, Alex Morison, Michael Smith and Jordan Haggit

June 16, 2021

Responses to overload as foundation for a resilient project and organizational health metric

Marisa R. G. BIGELOW, Lisa J. DOUGLAS, Alex MORISON, Michael W. SMITH, and Jordan M. HAGGIT
Mile Two, LLC, Dayton, OH, USA

ABSTRACT
It can be difficult to understand changes to team performance in the moment and at a larger organizational scale, especially when adapting to unique evolving pressures and opportunities. In some teams, situational awareness with respect to four responses to saturation - shedding load, reducing thoroughness, recruiting resources, and shifting work in time - may not be formally measured, or not measured collectively. To translate theory into practice and evaluate resilience at a project level, we are measuring the four responses to saturation on a regular cadence using a new project monitoring representation for communication and comparison at all organizational levels. Part of a larger project health metric, preliminary data for the four responses to saturation indicate an anecdotally significant increase in communication and awareness within the teams.

KEYWORDS
Situation awareness; engineering; saturation; metrics; project health; organization

INTRODUCTION

Most metrics used at the highest levels of an organization are often disconnected from the actual work being done and are narrowly defined to report on one aspect of a particular department or group within a much larger context. We are exploring a Project Health Measure that aggregates 6 metrics at the project level, enabling all the stakeholders at and above this level to have a shared understanding of the current state of the project as well as the ability to drill into each dimension that impacts the score. The Project Health Measure includes 1) team interactions, 2) artifacts produced by teams, 3) the quality of the software being developed, 4) how the project is progressing compared to the project plan, 5) a degree of customer satisfaction, and 6) a measure of project resilience. In particular for this context, project resilience uses the four dimensions defined by Woods and Hollnagel (2006) to capture the responses to saturation or overload: shedding load, reducing thoroughness, recruiting resources, and shifting work in time. With finite resources and evolving constraints, project teams and organizations are constantly navigating a complex trade-off space. A system's ability to respond to new demands can become saturated when workload bottlenecks exhaust the system's capacity to maneuver and adapt.

Organizations can greatly benefit from understanding how this construct can be measured, and resiliency increased, for the teams of people who do the work of the organization. Resilient teams are more effective and productive (Stoverink, Kirkman, Mistry & Rosen, 2020). However, the metrics used at most organizations are typically key performance indicators (KPIs) that reflect potential growth (e.g., qualified sales leads, employee engagement, NPS) but are by their nature abstracted and disconnected from the actual work being done. Project managements and software development teams also have KPIs to track things like bugs, budget, velocity, and volatility but these measures are typically focused on building the software and neglect some of the more intangible interactions that happen at the team level, like how people are getting along, is the customer happy with pace and progress, is the team producing what it says it will, or, importantly, how well the team is responding to changes thrown its way. It is also clear from searching the literature that organizational level metrics are usually employed when attempting to measure team level performance (e.g., Power & Waddell, 2004).

PILOT STUDY

We are currently piloting a subjective project health measure (referred to as Project Pulse) that aggregates six metrics at the project level from periodic survey measurements. The goal is to trace the critical decision points during the project in which the project lead and technical leads respond to increases in workload and significant changes in client direction. Project leads will evaluate the potential benefits and risks to document and support the decision of which class of response their team will follow.

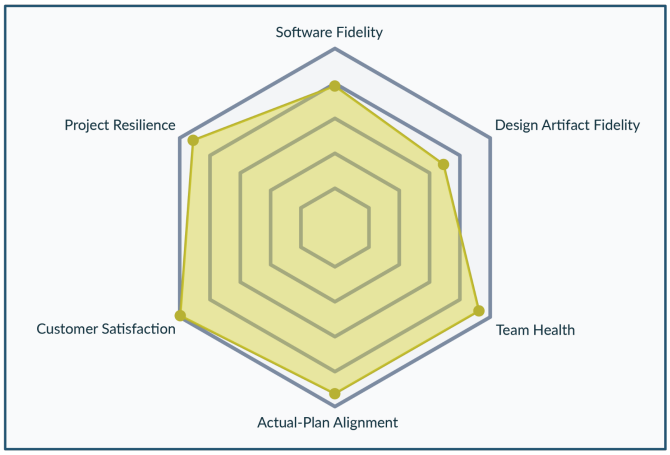

Figure 1. The radar graph for the Project Health Measure.

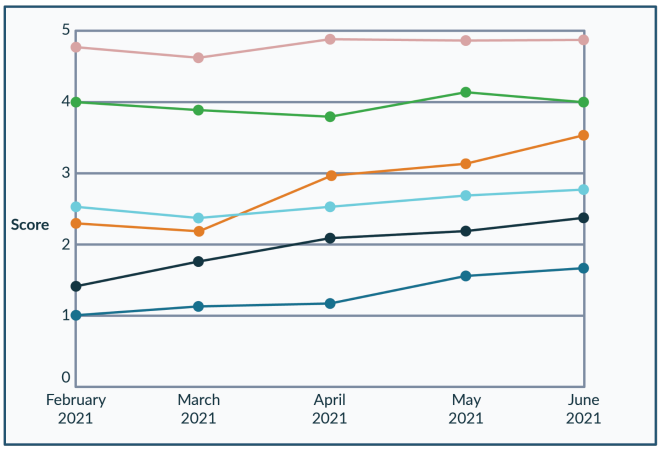

Figure 2. The time series for all six dimensions of the Project Health Measure.

Project Health Measure

Project Pulse measures six dimensions: 1) Team Health - team interactions, 2) Design Artifact Fidelity - research, design, and software artifacts produced by teams, 3) Software Fidelity - the quality of the software being developed, 4) Actual-Plan Alignment - how the project is progressing compared to the project plan, 5) Customer Satisfaction - customer or client contentment, and 6) Project Resilience - a measure of various responses to overload or saturation. The dimensions are represented as both a radar chart (Figure 1) for the latest measurement period and as a time series graph (Figure 2) for tracking changes over time. To support the project, we assembled a small team to build an online interactive application. However, the focus of this paper is narrowed to the measure of project resilience—four responses to saturation and overload described below. Specifically, the goal of the resilience metric within the Project Pulse pilot is to facilitate identification and understanding about the critical decision points during a project regarding the four responses. Would, or does, this approach to measurement via Project Pulse help a project lead better understand and communicate the current state of resilience? How does this awareness impact decision-making across the organization?

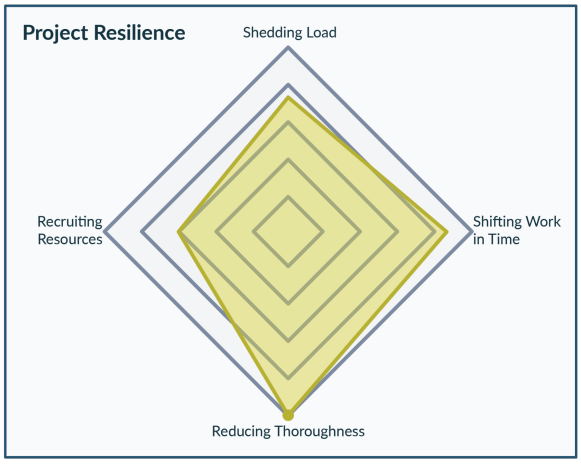

Figure 3. The radar graph for the Project Resilience measures.

Project Resilience

The four responses to overload (Figure 3) are coping mechanisms seen across various domains such as healthcare, aviation, etc. (Woods & Hollnagel, 2006). These categories were selected for the health metric as they can be tractably contextualized in the realm of software product development. For instance, shedding load refers to deprioritizing work, such as dropping tasks or requests. Reducing thoroughness reduces the amount of time allocated to complete the work, which could mean accelerating the pace of artifact generation relative to the team's underdeveloped work domain and problem understanding. Recruiting resources means to add people, capital, time, or material goods to alleviate workload pressure. Shifting work in time extends or delays the period in which work is completed, which could be exemplified as pushing tasks to the next sprint in an agile software development workflow. The four responses to saturation are further divided into tactical and strategic levels, where shedding load and reducing thoroughness are tactical responses useful in the short-term so long as priorities are well understood, and recruiting resources and shifting work in time are strategic with a longer-term focus. All four responses together give an awareness of the team's tendencies and explicit trade-off decisions.

Methodology

Our initial plan was to have project leads and technical leads report changes to workload and client direction by evaluating the potential benefits and risks to document and support the decision of which class of response their team will follow and subjectively reporting out biweekly. Table 1 shows an example of the resilience metrics questions from the biweekly survey. However, it was noticed for the first reporting period that the several project leads intended to include the entire collaboration team (research, design, and development) in the Project Pulse assessment. We are piloting with more consistent projects that we hypothesize would have similar tempos. However, if the pace and tempo dictate a different sensitivity, reporting also could be done ad-hoc. During and after the pilot period is completed, we will also query the project and technical leads about how they framed their decision-making for future, current, and past junctions. The aspect of risk is important to include when considering the consequences of actions taken by the project and technical leads. For instance, sacrifice decisions will inevitably arise with conflicting goals as in the case of multiple stakeholders or a shift in clients. Priorities must be balanced within limited resources, though it may not be explicitly apparent what opportunities the team may be giving up if only limited to considering tactical solutions. An affordance to articulate risk at the assessment time is provided by a free form text box in the Project Pulse application.

Table 1. The Project Resilience metric questions

Response to Saturation	Question
Recruiting resources	Did the team attempt to add resources of any type to the project (people, materials, capital)?
Shifting work in time	Did the team or any team member have to change any deadlines for any reason?
Shedding load	Did you or any team member need to reduce workload for any reason?
Reducing thoroughness	Did you, the team, or any team member reduce the thoroughness of your work (e.g., shortcuts, shortened deadlines) to improve efficiency?

We are currently evaluating two ways to use a Likert scale to measure the resilience responses. Using the example of *shed load,* one option is to define the scale from 1 to 5 where 1 means no load has been shed during the evaluation period (e.g., a week) and 5 means a lot of load has been shed. Alternatively, the *shed load* could be measured using a relative change where 1 means there is no change compared to the previous evaluation period and 5 means a lot of change has occurred. In an effort to understand which measure might be more sensitive in our environment, some project leads see their previous data entries when they fill out the questionnaire while

others see a blank questionnaire every time data is submitted. Early data suggests it is more efficient for the project leads to see previous data so they can quickly evaluate whether something has changed since the previous reporting period.

Besides the individual dimensions' scores, project leads have two indicators they can switch that will communicate their current stance on project risk to upper leadership: action taken and support needed. Each question has both options to show if a particular aspect in one of the dimensions is at risk and if action has already been undertaken to address it. Initial data indicates a clear need for more communication around both options because the current context is unclear, particularly from question to question if the meaning remains consistent.

Additional perspectives from the team (e.g., quality assurance, technical leads, machine learning engineers, designers, software engineers) will weigh in on the decisions and provide contrasting views of the problem. For instance, task management tools for development and quality assurance such as JIRA, Gitlab, and Monday.com have specific metrics that act as a proxy for different types of responses. Changes in story point estimates and task deadlines can signal purposeful shifts in work further out in time by providing more resources and duration to execute the work. Adding contributors to a task indicates the team has recruited resources for that task. These actions in isolation, however, are not sufficient evidence for a particular type of tactical or strategic response. The action must be contextualized with the contributing factors and environmental conditions surrounding the decision through deeper feedback.

Unintended Consequences

Several unintended consequences were identified just before and at the pilot launch in the initial Knowledge Elicitation (KE) sessions with project leaders and other stakeholders. KE refers generally to interview techniques in the flavour of Flanagan, Klein, and others (e.g., 1954, 1989). First, we identified one question in the resilience dimension that could be problematic. If a responder wants to maintain the appearance that their work is of the highest calibre regardless of extenuating circumstances, the question, "Did you, the team, or any team member reduce the thoroughness of your work (e.g., shortcuts, shortened deadlines) to improve efficiency?" (Table 1. Question 4) might be at risk for a less than honest response. The culture at Mile Two–a highly collaborative systems and learning-forward culture–largely mitigates this risk but it should be noted that in many organizations this type of culture is not present and may increase the need for creating appearances. It is worth exploring how this question, as well as the other questions in the resilience dimension, may be related to the Efficiency-Thoroughness Trade-Off (ETTO principle) described by Erik Hollnagel (2009). His assertion that sacrificing thoroughness for efficiency is normal and we would benefit from understanding in what situations that is likely to happen in our own organization. Additionally, there could be privacy concerns both for this question and question 3 in Table 1. A responder may be hesitant to answer honestly because they may feel they are identifying themselves as the one who is underperforming, especially from an outside perspective. They were performing to the best of their abilities, of course, given the imposed pressures and constraints.

Second, across the questionnaire there is validity risk at the item (question) level within each dimension, including resilience. We have diverse project leads, projects, internal discussions, and how they interpret each question and make decisions about how to respond requires more research. We already know that some project leads find some of the scales (intended to be 1 = worst and 5 = best for all questions) ambiguous. One project lead mentions they will respond 3 because they think things are going well. Other project leads will respond with a 5 as they consider the project going as expected. One of the ultimate goals of Project Pulse is to understand performance across the organization and misalignment at the individual question level will negate this goal.

Third, we have found in early analysis that the word "resilience" does not resonate with most people in the same way it does with academics or the consumers of papers such as this. Here again, it is important to present the questionnaire in terms that are easily interpretable by the responder. In-depth research will help us create better questions and result both in better validity and reliability.

NEXT STEPS

One long-term goal is to include additional information or signals that can easily communicate risk - an indicator of whether the measured change over time is actually a risk. This relates to changes in the time series where the graph shows an upward or downward trend. For example, there are some times when shedding load is expected and not a risk to the overall health of the project; at other times this action may indicate a risk and should be communicated as such. The differentiation can enrich the communication and conversation about the project. Interview data are clear that consumers of the Project Pulse app need more communication around the *action taken* and *support needed* indicators. A longer-term goal will be to add functionality to include space for elaborating on the selected indicator, especially to help contextualize the responses relative to the risk identified.

The Project Health Measure can support anticipatory risk mitigation in the future by cataloguing various project paths and surfacing characteristics of the trade-off spaces. As mentioned earlier, the four responses to saturation can be separated into tactical and strategic categories. Although any of the four could be appropriate depending on the circumstances, the distinction between tactical and strategic highlights a fundamental trade-off between

acute and chronic risks (Hoffman & Woods, 2011). Proactive strategies to manage these risks can help explicitly navigate the underlying Resilience-Optimality Trade-off that underpins the motivation behind the Project Resilience measure. Project management particularly struggles between more efficient, optimal planning and the need to be flexible, adaptive to changing dynamics. The various examples gathered through this effort can tangibly illustrate how diverse teams reasonably managed evolving trade-offs in retrospect.

IMPLICATIONS

The project metrics inform the next level higher in the organization at the portfolio level, in the pilot's case, which acts as a baseline for gauging the overall portfolio health. Traditionally, the portfolio directors relied heavily on the program managers and project leads' reflections on the various teams' progress. The health metrics would provide a similar point of comparison across the projects in collaboration with the leads' reflections. The evidence would be contextualized by each unique circumstance, but also generalized using the resilience responses to overload framework. The impact goes further to the executive level by then comparing across portfolios using the same basis at the micro-team level to abstract up to the macro-portfolio scale. The framework becomes a common language connecting the sharp and blunt ends of the organization, as well as demonstrating mindful team progress to other stakeholders (i.e., customers, operators).

The resilience metric applied to team project work would be of benefit to any organization that performs ongoing work using dedicated teams. The current test case focuses on a software development consultancy, though the work can be generalized to other domains. Organizations with multiple, distributed layers and teams could benefit from this work by applying the responses to saturation framework as a comparison metric. The metrics can enable teams and organizations to learn from previous project experiences to navigate future complications more effectively and create a larger learning culture. Locally rational decisions are contextualized in the flow of normal project work and thus, can be generalized using a universal framework for retrospective analysis. A corpus of project cases can inform future decision making by comparing similar circumstances and considerations. The different responses could also provide a prospective topology for alternative decision-making paths. Project leads, especially, can envision and simulate their responses to future pressures given the generic categories and relevant examples.

CONCLUSION

Work-level metrics can enable teams and organizations to learn from previous project experiences to navigate future complications more effectively and create a larger learning culture. Locally rational decisions are contextualized in the flow of normal project work and thus, can be generalized using a universal framework for retrospective analysis. For technology development organizations in particular, project and program management drive the product life cycles. Measuring resilience at the project level can provide important insight into team performance, identify issues and risk sooner, and highlight relationships between different layers of the organization and stakeholders. A corpus of project cases can inform future decision making by comparing similar circumstances and considerations. The framework and visualizations are in development now for contextualizing the trade-off space and comparing team performance around critical junction points.

ACKNOWLEDGEMENTS

We want to thank the Mile Two management for supporting us and for providing feedback in the development of this work. We also thank the participants of the pilot study as we continue to gather data and make improvements.

REFERENCES

Flanagan, J. C. (1954). The critical incident technique. *Psychological bulletin, 51*(4), 327.

Hoffman, R. R., & Woods, D. D. (2011). Simon's slice: Five fundamental tradeoffs that bound the performance of human work systems. In 10th International Conference on Naturalistic Decision Making, Orlando FL (pp. 5-31).

Hollnagel, E. (2009). *ETTO Principle: Efficiency-thoroughness Trade-off: Why Things That Go Right Sometimes Go Wrong.* Boca Raton, FL: CRC Press, Taylor & Francis Group.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on systems, man, and cybernetics, 19*(3), 462-472.

Power, J. & Waddell, D. (2004). The link between self-managed work teams and learning organizations using performance indicators. *The Learning Organization*, *11*(3), 244-259.

Stoverink, A. C., Kirkman, B. L., Mistry, S., & Rosen, B. (2020). Bouncing back together: Toward a theoretical model of work team resilience. *Academy of Management Review, 45*(2), 395-422.

Woods, D. D. & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton: CRC/Taylor & Francis.