Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

EVOLVING A GROUNDED APPROACH TO BEHAVIORAL COMPOSITION

Mayuri Duggirala
Mukul Malik
Suman Kumar
Harshal G. Hayatnagarkar
Vivek Balaraman

Human Centric Systems
TCS Research
Pune, 411013, INDIA

ABSTRACT

Human behavior simulation models having cognitive, physiological and social dimensions of behavior has been of interest for a long time. Within this landscape however work has lagged in 2 vital areas. There has not been much attention paid to models grounded in the behavioral sciences nor on techniques to create such models. In this work, we address both lacunae. We present a compositional approach to create grounded human behavioral simulation models and demonstrate its use with an example. The compositional approach has 3 components, a behavioral science element, a statistical and a computational. The behavioral component uses approaches to theory development in the behavioral sciences particularly in the realm of organizational behavior. The statistical element derives valid chains of relations. Finally the computational element converts the base model into a simulate model. We demonstrate our approach by creating and executing a simulation model in the workplace domain.

1 INTRODUCTION

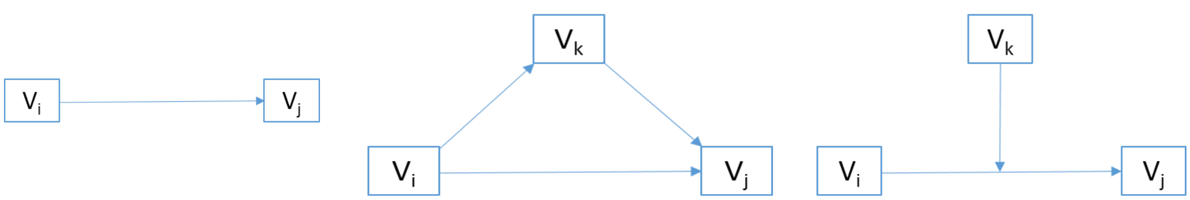
The creation of realistic human behavioral simulation models having cognitive, affective, physiological, and social facets of behavior has been of interest for a long time. Silverman (2004) introduced an approach to model human performance based on literature from behavioral sciences whereas Hudlicka (2003) shared a cognitive-affective architecture where human behavior would manifest from various profiles. In our work (Singh et al. 2016), we detailed an approach to compose models which we called *fine-grained* as it incorporates detailed behavioral relationships, and *grounded* for sourcing these relationships from either field studies or peer-reviewed literature. In this paper, we extend our earlier work and show how a multidisciplinary and integrative lens is required to compose even simple models of behavior, and share guidelines and criteria from each of the participating disciplines. We then demonstrate this approach with a composed model which simulates a support services organizational environment. Theory building / development is integral to the progress of any scientific endeavor including the behavioral sciences. Newer theories have been built on the insights of the past, wherein existing theory has been integrated or synthesized with newer behavioral dimensions or insights. Theory has been described as an explanation of processes or events (DiMaggio 1995); as a set of assertions that identifies important variables, while outlining why they are important, and as a specification of the conditions under which these assertions should be related (or not) (Campbell 1990). We view our work as an exercise in behavioral science theory integration, synthesis and development coupled with statistical measures which deal with extending the scope of individual relations plus validity of the overall composition and a final computational harness to create simulate ready behavior models. Our main contributions are: an approach for composing models of behavior founded on theory integration within the behavioral sciences, a statistical method based on meta-analysis to get generalized forms of individual relations and measures for the entire composition, articulation of computational aspects to convert the base behavioral model created through the preceding two steps into a simulate ready model and finally a demonstration of the use of our compositional approach to compose a simulation model in the workplace domain. The rest of the paper is organized as follows: we first discuss composition and its challenges from the perspectives of behavioral science, the statistical and the computational. Next, we propose an integrated approach to composition and present requirements for composition. Finally, we present a demonstration of our approach along with a discussion of our findings and suggestions for future research.

2 OUR APPROACH TOWARDS COMPOSITION AND ITS CHALLENGES

Given a set of user requirements of a model, *composition* is the systematic process of selecting and assembling a theoretically and empirically valid and unified model from a collection of reusable components. Model composition has been explored in the general context (Petty et al. 2003) as well as in the simulation context (Petty et al. 2014). In this section, we begin by describing forms of composition in the space of human behavior models, some of the challenges associated with composition, and go on to discuss how the behavioral sciences, statistics, and computer science need to address them.

2.1 Composition

Let's assume that $V = \{v_1, v_2, \dots v_n\}$ is a set of behavior variables. These behavior variables are connected to each other via a set of relations $R = \{r_1, r_2, \dots r_m\}$, where $r_i; i = 1,2, \dots, m$ describes a relation between two or more variables in V. Each relation can take one of the forms described in Figure 1. In its simplest form, a bivariate relation (Fig 1-a) captures an association between an independent or predictor variable (v_i), and a dependent or outcome variable (v_j). Apart from bivariate relations (Fig 1.a), there are two other relation forms within the behavioral sciences incorporating two forms of explanatory variables. Explanatory variables, (variables which explain the relationship between two or more predictor and outcome variables) are known to take two forms: mediation (Fig 1.b) and moderation (Fig 1.c Baron and Kenny,1986). Mediation is said to be in effect when *a variable explains the relationship* between the predictor and outcome variable. Moderation exists when a variable *influences the strength* of the relationship between a predictor variable and an outcome variable. Please see Hayes (2013) for an detailed discussion.



(a) Bivariate relation (b) Mediation (c) Moderation

Figure 1: Different types of behavior relations(adapted from Hayes 2013).

Thus, composition is a process to assemble a set of relations in R, such that each relation includes one or more behavioral variables, belongs to one of the three types of relations, and participates in composing a longer chain. This longer chain is a new hypothesis which incorporates and integrates the insights found in the relations. So if v1→v2 and v2→v3 are two distinct relations obtained from the set R then a possible composition is v1→v2→v3. We call v2, which occurs in both relations as a *link variable*. Link variables can denote relationships of several kinds: mediators, moderators or association. The underlying theoretical framework for composition will guide this decision.

Note that we would require common linking variable(s) across studies (in the above case, v2). Unconnected relations in the form of v1→v2 and v3→v4 cannot be considered as suitable for composition unless we can justify that v2 and v3 are semantically and statistically similar if not identical. More detailed considerations for our composition approach are presented in sections 2.3, 2.4 and 2.5 from perspectives from organizational behavior, statistics and computational modeling, respectively.

2.2 Challenges

In composing fine-grained and grounded human behavior models, even simpler forms like chain of relationships pose challenges because multiple disciplines need to agree upon the process of composition and its outcome. Behavioral sciences such as psychology, management and organizational behavior conduct studies to test behavioral hypotheses proposing relationships among specific variables of interest. Each such empirical study reports insights valid within its context. Our composition process relies on a corpus of these hypotheses and their results from different studies to develop a composed model.

One challenge therefore from a behavioral science standpoint derives from the fact that definitions of the behavioral factors in one study are likely to vary across studies due to different theoretical operationalization, empirical measurements, treatments of the variables or due to differing study contexts, to name a few. The opposite problem is where two variables have different names but are semantically the same. This is called the *Jingle-Jangle problem* in the social sciences literature (Larsen et al. 2013). These considerations are particularly important when the variables concerned are link variables.

For example, a behavioral hypothesis could be in the form of the following statement: Teaching style has a positive impact on an individual's motivation to learn (*Teaching style*→*Motivation*), while another hypothesis could suggest that motivation to learn is positively linked to mastery (*Motivation*→*Mastery*). Here we could treat *Motivation* as the common or linking variable and create a new composed relation or hypothesis as follows: *Teaching style*→*Motivation*→*Mastery*. Here we have the variable *Motivation* common between the two hypotheses, allowing us to create the extended relation linking *Teaching style* as an indirect predictor of *Mastery* although this hypothesis is not originally theorized or tested. In creating this extended relation, we need to ensure that the definitions of *Motivation* in both individual relation fragments pertain to the same underlying explanation of *Motivation*, either *Intrinsic Motivation*, *Extrinsic Motivation*, or other forms of *Motivation*. Having thus concluded that the underlying construct of *Motivation* is same across the two relation fragments, we are then in a position to create a longer composed chain: *Teaching style*→*Motivation*→*Mastery*. This process requires domain knowledge in the behavioral sciences and an underlying behavior relations repository which is structured around an ontology of constructs in the behavioral sciences. We have made initial efforts in creating such an ontology (Duggirala et al. 2015) . The above example is illustrative of such scenarios in past research where we encounter examples of hypotheses which have not yet been connected, but together offer a richer understanding of an outcome of interest or an extension of existing theory.

Since composition refers to the synthesis or combining of *different* variables and their relationships across studies, there is a challenge in ensuring that apart from ensuring semantic equivalence of link variables they also have similar statistical properties, in terms of the distribution of the variable (mean and standard deviation), and similarity in measurement scales for the variables.

There are also computational challenges in converting a composed behavioral model to a simulation ready state. It is thus imperative to understand implications of this compositional approach from the view points of the behavioral sciences, statistics, and the computational sciences. In the rest of this section, we discuss each of these viewpoints as well as present an integrative perspective.

2.3 Composition from a Behavioral Sciences Viewpoint

With respect to behavioral science there are three key challenges. The first challenge is that of situating this work in the larger realm of how knowledge growth takes place in the behavioral sciences. The second is that since each element in our model is using a relation from a specific study, we need to be able to generalize from this study to a larger population size. The third is the *jingle-jangle* problem discussed in section 2.2. In this section we discuss how we deal with each of these.

With regard to theory development in the behavioral sciences, researchers often build on existing constructs in several ways, such as by discovering new relations among existing hypotheses, or by synthesizing/integrating similar relevant dimensions into a single coherent theoretical framework with a common underlying theoretical logic or rationale. For example, in a recent study on personal identification (Ashforth et al. 2016) the authors define personal identification, distinguish it from other forms of identification and identify pathways by which individual identities are constructed. That paper (and similar review papers) summarize past research and evolve a richer theory of behavior incorporating related but distinct dimensions which serve to further our understanding of the behavior being studied. Similarly, in another study, (Song et al. 2016) the authors evaluate the effect of selfishness and selflessness combined with emotions of sympathy and vying in their impact on pedestrian evacuation. These examples suggest how authors have attempted to combine disparate but related behavioral variables to examine their joint impact on outcomes of interest.

The next challenge we encounter in the behavioral sciences while evolving an approach to composition is that of generalizability of insights from a study or set of studies to a larger population. Most studies in the behavioral sciences are done on specific samples testing a few study hypotheses. The meta-analytic paradigm within the behavioral and social sciences derives an effect size by synthesizing results for similar or identical behavioral variables across samples, using two primary approaches: fixed effects models and random effects models. An effect size refers to the estimate of the effect of the behavioral variable(s) of interest over the larger population (as opposed to samples within individual studies). Here, fixed effects models assume that studies in the meta-analysis are sampled from a population whose underlying average effect size is fixed, thus resulting in homogenous sample effect sizes (Hunter and Schmidt 2000). On the other hand random effects models assume that population effect sizes vary randomly and are therefore heterogeneous (Hedges 1992). A large proportion of meta-analytic research uses fixed effects models as compared to random effects models while the latter have been found to be more appropriate for real world data (Field 2003). In line with this suggestion, since our approach rests on relations derived from studies using real world datasets incorporating greater heterogeneity and generalizability, the random effect model is more appropriate.

The third challenge is the jingle jangle problem. In order to address the problem, we currently use inputs from a domain expert in behavioral science to both identify semantically similar but differently named constructs and to distinguish between semantics of the same variable name used in different studies. In future we would like to address this through techniques such as latent semantic analysis for identifying semantically similar terms (synonymy problem) as well as using textual disambiguation techniques to address the polysemy problem.

2.4 Composition from a Statistical (Empirical) Viewpoint

Composition as we have discussed, involves linking together a set of relations. For each pair or tuple of variables there may be multiple past studies in the corpus which have determined a relation between these variables. The first statistical requirement is therefore in selecting a relation is that the study findings must indicate that the relationship between those variables is strong. The second requirement is that since we will be using each relation with its associated tuple of variables in a simulation, we must be able to determine the change in one variable (usually the dependent variable) given a change in the other variables (independent variables). And finally, we must be able to use these selected relations to get a composed model that can be then taken up for further processing by the downstream computational element.

In most quantitative empirical studies regression and correlation analysis are both used to study linear relationships between variables. Correlation is the measure of the strength of the linear relationship between two variables. Therefore we can use the correlation measure to satisfy the first requirement on the principle that only variables which are significantly correlated can be used as a behavioral fragment in composition. Meanwhile regression, specifically linear regression is used to predict the value of one variable for a change in the other (by using the *line of best fit*). We therefore use the linear regression equation for the second requirement. A relationship between these variables can be obtained in the form of a linear regression, if we have the correlation as well as mean and standard deviation of each of these variables. Generally linear regression equation take the form $Y = \beta_0 + \beta_1 x + \varepsilon$, Where β_0: y-intercept; β_1: slope of the line; ε: error term. From Rodgers (1988), the slope β_1 and intercept β_0 term in regression can be estimated as follows (please see appendix A for more details):

$$\widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{x} \quad , \quad \widehat{\beta_1} = r_{xy} * \frac{Sy}{Sx} \,. \tag{1}$$

Next, different behavioral fragments (in the form of linear regressions) need to be connected together in order to get a composed quantitative model. For example, consider three distinct behavioral fragments in the form of v1→v2, v2→v3 and v3→v4 which we want to link as v1→v2→v3→v4. One major issue for this composition, is that the reported mean and standard deviation of behavioral variables in different individual studies are not same. In order to get estimates of regression beta coefficients from above equation (1), one needs to have a common value of mean and standard deviation of variables. Our approach to solve this issue is get a pooled mean & pooled standard deviation of common behavior variables. Their pooled mean and standard deviation is measured using the following equations:

$$\overline{\mu}_y = \frac{n_1 m_1 + n_2 m_2 + \cdots + n_k m_k}{n_1 + n_2 + \cdots + n_k} \quad ; \quad S_y = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + \cdots + (n_k-1)S_k^2}{n_1 + n_2 + \cdots + n_k - k}}$$

Where $m_1, m_2, \ldots, m_k$; $s_1, s_2, \ldots, s_k$ and $n_1, n_2, \ldots, n_k$ are the means, standard deviations and sample sizes of the fragments from individual studies.

2.5 Composition from a Computation (Simulation Modeling) Viewpoint

The behavioral perspective and the statistical together have helped us to obtain a valid, quantifiable composed model. We need the computational element to go from this to a simulate-ready model. A human behavior model cannot be simulated in isolation, and needs to be immersed in a context usually specified by environmental and process models. Our simulation model therefore consists of three interconnected aspects: environmental, process and behavioral.

The environmental aspect consists of descriptions of the world in which the agents are situated and with whom the agents will interact. Different scenarios, events and associated intervention strategies will also come under the environmental dimension. The process aspect is particularly important while modelling contexts where human entities work within a certain rhythm specified by a process as for example, in case of a support services organization, where the process model has a specification of how work is received every day, how it is distributed to available team members, how performance of team is identified and how metrics of interests are reported. Finally, the behavioral aspect captures the model of individual agents. Environment and process structures and mechanisms may lead to events which triggers appropriate behavior. Coordination between the environmental, process and behavioral can be seen as the composition challenge from the computation point of view. The principal considerations are:

Temporal granularity of variables**:** Some variables are slow varying and others are not. For example, stress of an individual may change every hour, while the team backlog may not. As these variables are connected/linked to each other in the simulation model, temporal variability of each of them needs to be taken into account. A convenient approach to overcome this problem is to bound variables into different time zones. For that, specific timescale events can be specified, which allows variables to vary appropriately.

Computable models**:** Computer simulations are based on computable models, such that changes in one or more variables (independent variables) affect one or more related variables (dependent variables).

Regression functions, and rules are commons forms of computable models which capture such relationships between independent variables and dependent variables.

2.6 Composition from an Integrated Viewpoint

2.6.1 Caveats

A few caveats on the composition process. The first of these is that we differentiate between a chaining composition and a mediation. A chaining composition may consist of a linked chain of individual relations where the dependent variable in one relationship is the independent variable in the next linked relation. However, this should not be confused with a mediated relationship which too appears on the surface as a composed relation, for example, v1→ v2 → v3. However, this is a single construct and cannot be broken into say 2 relations such as v1→v2 or v2→v3. A similar consideration would apply to a moderation hypothesis.

Another consideration, this from a statistical standpoint, is that a direct relationship between a dependent and independent variable can only be considered for linking in a composition if they are highly correlated because higher correlation leads to higher coefficient of determination R^2. However, several challenges emerge in this process, such as when study variables do not show very high correlation because of lack of sample size and/or when the study itself is only theoretical, and not supported by an actual empirical study. To overcome this problem, one may choose to link a set of variables) which are highly correlated to both input and outcome variable giving rise to more explanatory power. At the same time, when one increases the length of a chain, this also increases the sum of squares for the model. Hence one should optimally decide the number of explanatory variables.

2.6.2 Guidelines for composition

Based on the discussions so far, we give the guidelines for composition (outlined in Figure 2):

- **a.** Select semantically similar variables and theoretically complementary hypotheses in all of the studies used in the composition.
- **b.** Select common variables that can be linked in a theoretically and statistically valid manner across relations from R.
- **c.** Select variables with higher correlations to both input and output variables balancing the error sum of square for the model.
- **d.** When common mean and standard deviation of variables are not available for different meta-analytic studies, individual mean and standard deviation for behavioral variables are used.
- **e.** Effect size (correlation coefficient in our case) is used to calculate regression coefficients for composing the model.
- **f.** In our current approach, we use effect size estimates only based on random effects models to achieve greater generalizability of effects to a larger population.
- **g.** Use of these effect size estimates in composing behavioral fragments into a longer chain.
- **h.** Running simulation experiments based on the composed behavioral model.

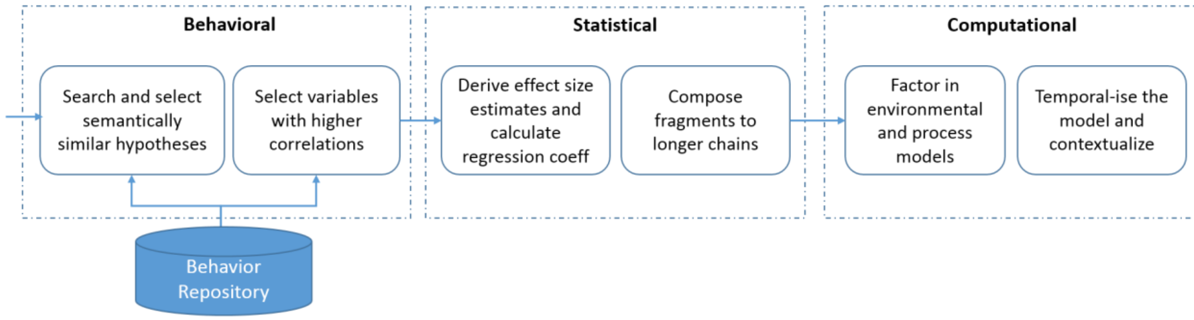


Figure 2: Behavior composition process.

3 APPROACH DEMONSTRATION, EXPERIMENTS AND RESULTS

In this section, we demonstrate our approach to compose a simulation model of a support services organization, shown above in Figure 2. Now, we would like to define terms we will be using in the remaining section pertaining to functional, operational parameters of support services organizations as well as behavioral dimensions.

- Task arrival rate: Rate at which new tasks arrive at the beginning of each work day
- Negative Affect: Negative emotion felt by an employee
- Work stress: Work-related stress felt by an employee
- Backlog: Unfinished tasks for an employee at the end of each work day
- Performance: Capacity of an employee to do work related tasks per hour
- Turn-around-time: Time duration in hours for a task from arrival to completion
- Bench strength: A pool of employees that is assigned work in case of absenteeism or increased workload, and calculated as a percentage of original team size

3.1 Goals of our model, variables and relations

The goal of our model is to understand how workload and workload related stress can impact performance both of the individual in a team and the team itself. The variables of interest are workload, negative affect, stress and job performance. As discussed in earlier sections, we have three relations: Workload → Negative affect, Negative affect → Stress, Stress → Job Performance. The cumulative job performance then impacts workload (for example if job performance has been poor, it will mean that apart from daily workload, we also have to deal with backlog).

3.2 Behavioral Science Perspective

The considerations for composition outlined in section 2.3 were applied to the selection as well as linking of the relations of interest. First, the aim of our work was to study the impact of workload and work related stress on performance. For this purpose, the research in these areas was surveyed extensively by a researcher in organizational behavior and the key variables of interest were identified: workload, negative affect, work stress and performance. The hypotheses which link these variables in past research were also identified. Past research shows links between workload→negative affect (Ilies et al. 2007), negative affect→stress (Hyun et al. 2017) and stress→performance (Beehr 2014). Based on the above review of research, the underlying rationale for linking these individual relations was defined such that the relationship between workload→negative affect→work stress→performance was clear. The composed model describes how the quantum of work influences negative experience related to work, in turn increasing stress at work which finally impacts performance. Thus individual relations studied in previous research were linked to form an integrated composed chain with negative affect as the link variable denoting association.

Many of the studies reviewed had published empirical findings on the above hypotheses based on individual studies in specific study contexts. As discussed in section 2.3, we aimed to achieve a more generalizable effect size for each of the above relations participating in the composition. Thus we used their meta-analytic effect size estimates in composing our model. These effect size estimates were based on our search of the Metabus repository (Bosco et al. 2015) which published rBUS estimates for each of the relations of interest. In the interest of greater generalizability of findings, random effects meta-analysis was used in our composition approach.

Much of the research reviewed to identify the relations of interest was characterized by several overlapping constructs. For example, in our literature review as well as meta-analytic search for studies on stress and performance, we took care to include only studies on work-related stress and performance and excluded other forms of these variables. It may be noted that the initial literature review for relations cited above spans workload, stress, negative affect and performance across different constructs and scenarios. For the final integrated composed model, however, only *work-related* constructs were identified and included. Similarly, there are several definitions of negative affect in the literature, such as 'distress', 'negative mood' and 'negative emotion'. After ensuring that these terms pertained to the same underlying experience of negative experience at work, the studies using these terms were included in the meta-analysis. Based on these considerations, we present our composed model in Figure 3. Table 1 has more details.

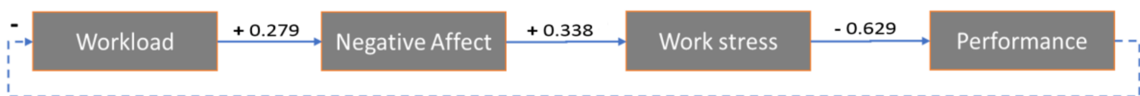


Figure 3: Human behavior model with beta coefficients indicating relationship strength.

3.3 Statistical Perspective

Once we have the theoretically composed model from the behavioral sciences, we need to convert this theoretical model into a quantifiable model that will serve as an input for simulation. For this purpose, we need to statistically integrate the distinct behavioral fragments of the composed model shown in Figure 3. We have used regression equations for integrating these behavioral fragments. Pooled means and standard deviations needed for computation of regression coefficients are calculated as described in section 2.4. Table 1 reports the pooled means and standard deviation and the regression equation for each fragment participating in the composition.

Table 1: Statistical measures required for composing behavior model.

Behavioral variable	**Pooled mean**	**Pooled standard deviation**	**Form**
Workload	3.651	1.581	Workload = Arriving tasks + backlog
Negative Affect	3.128	2.006	Negative Affect = 2.11 + 0.279 * Workload
Work Stress	2.379	1.421	Work Stress = 1.322 + 0.338 * Negative Affect
Job Performance	4.30	1.246	Job Performance = 9.63 – 0.629 * Work Stress

3.4 Computational Perspective

We simulate business-as-usual for a typical team working in a support services organization. The organizational *environment model* is detailed as follows: At the start of every business day, the team receives workload in form of tasks from the customer. Tasks are assumed to be independent and identical in nature. The tasks are gathered in a 'task pool', which is modeled as a synchronous queue. Once workload is arrived, the tasks are assigned to the available team members. Apart from the usual team a certain percentage of individuals are also available if extra workload arrives for a day. We refer this as bench strength. Project manager keeps track of the work completed and reports the metrics of interest to the customer. While simulating, different interventions such as spike in workload, high variations in workload, or crisis can be introduced. Project manager can apply different management strategies to cope up with situations, in case of any crises or other interventions. As shown in Figure 4, the *process model* covers the flow of tasks from their arrival to their completion. We assumed that task arrival pattern follows a normal distribution with mean of *1000* and standard deviation of *100*. The tasks are uniformly distributed to each individual in the team. On any typical day, an individual can spend 8 hours for regular work and additional maximum of 2 hours as overtime. The uncompleted tasks go back to the task pool as team backlog, which factor in the workload for next day. At the end of each day, team performance and other metrics of interests are reported.

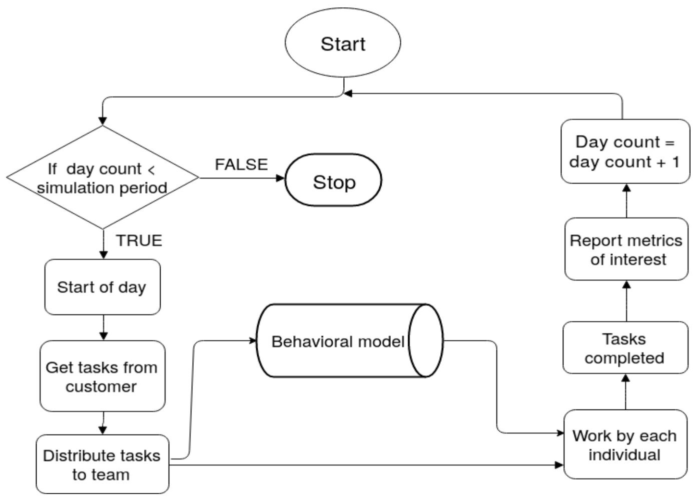


Figure 4: Process model for support services organization.

Behavior of each individual agent is a function of the behavior model, the current state of agent and the stimuli from the environment. Sections 3.2 and section 3.3 have led to the creation of a base behavior model. To generate a computable model for simulation, we take the base model and make it simulate-ready. As discussed in section 2.5, we need to resolve the challenge regarding the different temporality of behavior variables. In this context, we assumed that workload is changing on a daily basis, while negative affect, stress and job performance are varying on an hourly basis. As there is a mismatch in the temporality of variables, we introduced timescale events, namely per day and per hour. These events keep track of updating the variables appropriately in simulation. Once we have that simulate-ready model , we can further proceed with the experiments.

3.5 Experiments

In our experiment, we conducted two scenarios, for which team performance and average turn-around time (TAT) of tasks is monitored by varying bench strength. Each simulation result has been averaged over 10 simulation runs on the GAMA platform. In first scenario, we have the team working under the influence of workload and their performance is mitigated by amount of workload assigned to them individually. Each individual receives same amount of workload initially. As the simulation progresses, the distribution of tasks are based on the number of tasks remaining. Agents with lower workload are assigned more tasks. . Tasks of an absent agent are redistributed to other team members. In second scenario, we introduce a spike in work, by doubling the number of arrived tasks (2000) for a day. In this case, the implicit feedback loop of *performance* → *workload* plays a significant role in determining an individual's productivity. This can be seen in the visualization of team performance in both scenarios, and for different bench strengths. We first study variations in the average TAT for the simulated team. Figure below shows the average TAT in hours along with the presence and absence of a spike in workload. We observe that without work spike simulated team performs reasonably well as tasks gets completed within stipulated time. A bench strength of 6% is manages to keep TAT under control (TAT ~ 5 hours). As we introduce the work spike, TAT shoots up, as high workload has lowered the performance of team and tasks are waiting in a queue to be completed. The average TAT reduces to an accepted level (from 96 hours to 6 hours), only when the bench strength is increased from 0% to 15%.

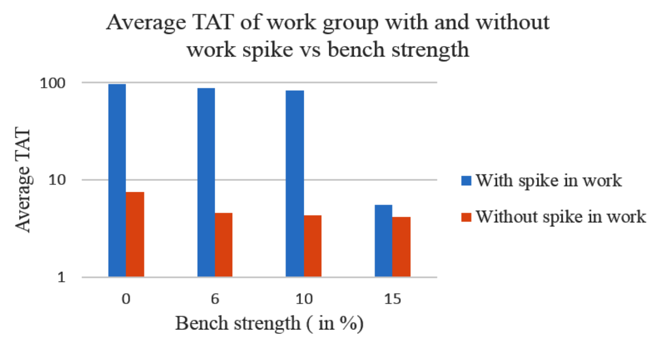


Figure 5: Change in average turn-around time.

In Figure 6, we have reported the overall performance of team with and without spike in workload. In absence of spike, the team feels less amount of stress, and hence less effect of bench strength. Thus we observe a high team performance in this case. However, varying the bench strength does slightly increase the team performance. In case of work spike, performance of team is significantly reduced as seen in the figure. As we increase the bench strength from 0% to 10%, workload remains a determining factor and performance of team suffers. Further increasing bench strength to 15%, however, overcomes the high workload, and team recovers from spike in workload. Therefore when bench strength is 15%, team performance is similar to the performance in no spike case. The bench strength of 15% allows risk free operation but most operational environments operate with lower bench strengths.

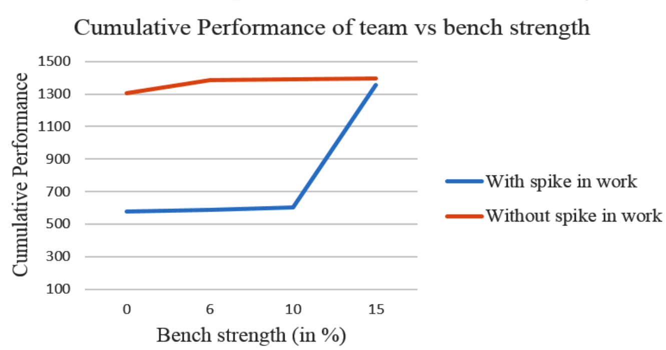


Figure 6: Change in team performance with/without work spike.

4 CONCLUSION AND FUTURE WORK

In this paper, we have extended our approach of composing fine-grained and grounded models of human behavior, with a focus on a multidisciplinary and integrative lens. The lens provides guidelines and requirements from each discipline to ensure that the composed model is consistent and valid from respective viewpoints. To demonstrate this approach, we have developed a composition model of a support services organization having variables of interests such as negative affect, stress and job performance and sourced their relationships from meta-analytic studies found in literature. In the present composed model, we have only used chaining of bivariate relations to demonstrate our approach to composition. Future work can extend this approach to incorporate mediation and moderation models as well.

A APPENDIX

Calculating Beta Values from Correlation Coefficients: Several simple algebraic forms for correlation can be used for computational purposes. A widely used form is as follows:

$$r = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum(X_i-\bar{X})^2}\sqrt{\sum(Y_i-\bar{Y})^2}}, , \quad \text{which can be further simplified to:}$$

$$r = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{(n-1)S_x S_y} \Rightarrow \sum(X_i-\bar{X})(Y_i-\bar{Y}) = r*(n-1)S_x S_y$$

One can obtain an unbiased estimate of regression coefficients by minimizing the error sum of square. After minimizing error sum of square, by using equation (1) the slope estimate of regression equation is:

$$\widehat{\beta_1} = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sum(X_i-\bar{X})^2} = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{(n-1)S_x^2} = \mathrm{r}_{xy} * \frac{S_y}{S_x}$$

Where S_x and S_y represent the standard deviation of the variables x and y respectively and r_{xy} is the coefficient of correlation between x and y. Hence, beta coefficients can be estimated using correlation coefficient via following relation, where, $\bar{Y}$ and $\bar{x}$ are means of Y and x respectively:

$$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1}\bar{x} \qquad \widehat{\beta_1} = \mathrm{r}_{xy} * \frac{Sy}{Sx}$$

REFERENCES

Ashforth, B.E., B.S. Schinoff, and K.M. Rogers. 2016. "“I Identify with Her,” “I Identify with Him”: Unpacking the Dynamics of Personal Identification in Organizations". *Academy of Management Review 41*(1):28-60.

Baron, R.M., and D.A. Kenny. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations". *Journal of Personality and Social Psychology 51*(6):1173.

Beehr, T. A. 2014. "Psychological Stress in the Workplace (Psychology Revivals)". Routledge.

Bosco, F.A., P. Steel, F.L. Oswald, K. Uggerslev, and J.G. Field. 2015. "Cloud-Based Meta-Analysis to Bridge Science and Practice: Welcome to Metabus". *Personnel Assessment and Decisions* 1(1).

Campbell, J.P. 1990. "The Role of Theory in Industrial and Organizational Psychology".

DiMaggio, P.J. 1995. "Comments on" What Theory is Not"." *Administrative Science Quarterly 40*(3):391-397.

Duggirala, M., V. Balaraman, A. Kumar, and D. Clarance 2015. "Building an Ontology for Behavioral Science." *Computational Social Science Summit 2015*, Northwestern University.

Field, A. P. 2003. "The Problems in Using Fixed-Effects Models of Meta-Analysis on Real-World Data". *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences 2*(2):105-124.

Hayes, A.F. 2013. "Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach". Guilford Press.

Hedges, L.V. 1992. “Meta-Analysis”. *Journal of Educational Statistics* 17(4):279-296.

Hudlicka, E. 2003. “Modeling Effects of Behavior Moderators on Performance: Evaluation of the Mamid Methodology and Architecture”. *Proceedings of BRIMS, 12*.

Hunter, J.E., and F.L. Schmidt. 2000. “Fixed effects vs. Random Effects Meta-Analysis Models: Implications for Cumulative Research Knowledge”. *International Journal of Selection and Assessment* 8(4):275-292.

Ilies, R., K. M. Schwind, D. T. Wagner, M. D. Johnson, D. S. DeRue, and D. R. Ilgen. 2007. “When Can Employees Have a Family Life? The Effects of Daily Workload and Affect on Work-Family Conflict and Social Behaviors at Home”. *Journal of Applied Psychology* 92(5):1368.

Hyun, J., M. J. Sliwinski, D. M. Almeida, J. M. Smyth, and S. B. Scott. 2017. “The Moderating Effects of Aging and Cognitive Abilities on the Association between Work Stress and Negative Affect”. *Aging & Mental Health* 1-8.

Petty, M.D., and E.W. Weisel. 2003. “A Composability Lexicon”. In *Proceedings of the Spring 2003 Simulation Interoperability Workshop* 2003:181-187.

Petty, M.D., J. Kim, S.E. Barbosa, and J.J. Pyun. 2014. “Software Frameworks for Model Composition”. *Modelling and Simulation in Engineering 2014*:4.

Rodgers, J. L., and W.A. Nicewander. 1988. “Thirteen Ways to Look at the Correlation Coefficient”. *The American Statistician*, 42(1):59-66.

Silverman, B.G. 2004. “Toward Realism in Human Performance Simulation”. *The Science and Simulation of Human Performance* 5:469-498.

Singh, M., M. Duggirala, H. Hayatnagarkar, S. Patel, and V. Balaraman. 2016. “Towards Fine Grained Human Behavior Simulation Models”. In *Proceedings of the 2016 Winter Simulation Conference,* edited by T.M.K Roeder, P.I.Frazier, R.Szechtman, E.Zhou, T.Huschka and S.E.Chick, 3452-3463. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Song, X., L. Ma, Y. Ma, C. Yang, and H. Ji. 2016. “Selfishness-and Selflessness-Based Models of Pedestrian Room Evacuation”. *Physica A: Statistical Mechanics and its Applications* 447:455-466.

AUTHOR BIOGRAPHIES

MAYURI DUGGIRALA, PhD (Management) from IIT Madras, works in behavioral science, particularly in organizational behavior research at HCS. Her email address is mayuri.duggirala@tcs.com.

MUKUL MALIK is a researcher at HCS working in the areas of Machine Learning and Data Mining, Regression analysis and Neural networks. He did his Masters in Statistics from IIT Kanpur. His email address is mukul.malik2@tcs.com.

SUMAN KUMAR is a researcher at HCS working in modelling & simulation and its real world applications. His email address is suman.kumar4@tcs.com.

HARSHAL G. HAYATNAGARKAR is a senior research associate with HCS, and a registered PhD candidate at Dept. of Computer Science at Virginia Tech. He is interested in agent-based simulations of large-scale social dynamics, especially for urban resilience. His email address is harshalh@vt.edu.

VIVEK BALARAMAN leads the HCS research group. He has worked extensively in case-based rule-based knowledge based representation and reasoning, experiential textual knowledge management, decision support systems, machine learning, natural language processing and in recent years in human behavior modeling and simulation. His email address is vivek.balaraman@tcs.com.