Uncorrected manuscript, accepted:
Birney, D.P., Beckmann, J.F., Beckmann, N., Double, K.S., & Whittingham, K. (in press). Moderators of learning and performance trajectories in microworld simulations: Too soon to give up on intellect!? *Intelligence*.

Moderators of learning and performance trajectories in microworld simulations: Too soon to give up on intellect!?

Damian P. Birney (University of Sydney, damian.birney@sydney.edu.au)

Jens F. Beckmann (Durham University, j.beckmann@durham.ac.uk)

Nadin Beckmann (Durham University, nadin.beckmann@durham.ac.uk)

Kit S. Double (University of Sydney, kit.double@sydney.edu.au)

Karen Whittingham (University of NSW, karen.whittingham@unsw.edu.au)

Corresponding Author:
Damian P. Birney
School of Psychology
University of Sydney
Sydney, NSW, 2006, Australia
E-mail: damian.birney@sydney.edu.au
Affiliations: School of Psychology, University of Sydney

Abstract

The burgeoning increase in the importance given to the role of non-cognitive factors in complex decisions making has led to calls to dethrone intelligence as the primary explanatory model of success. Using a combined experimental-differential paradigm and mixed-level modelling, features of 8 business microworld simulations were experimentally manipulated to investigate the incremental value of non-cognitive predictors of learning and performance trajectories. It was predicted that facilitating personality traits (openness and extraversion), growth/motivational mind sets (learning goals, need for cognition, and beliefs of malleability), and tentatively, emotional-regulation (managing and facilitating emotions) would moderate the impact of complexity and experience on performance and learning trajectories. Results based on a sample of 142 experienced mid-level senior managers suggest microworld simulations can be manipulated to be differentially sensitive to domain-specific differences in reasoning. Reasoning moderated learning effects and the capacity to deal with task complexity. Of the 16 non-cognitive factors investigated, only performance-goal orientations moderated performance trajectories over and above reasoning. These findings give reason to question the importance of non-cognitive dispositional and motivational factors in learning and problem-solving, over and above intellect. We argue that microworld simulations and mixed-level modelling approaches can support the experimental investigations needed for comprehensive, dynamic real-world training research.

Moderators of learning and performance trajectories in microworld simulations:

Too soon to give up on intellect!?

Damian P. Birney[1*]

Jens F. Beckmann[2]

Nadin Beckmann[2]

Kit S. Double[1]

Karen Whittingham[3]

Running Header: MICROWORLD PERFORMANCE MODERATORS

Keywords: microworld simulation, self-regulation; conative dispositions; cognitive ability; performance and learning

1 School of Psychology, University of Sydney, Australia
2 School of Education, Durham University, UK
3 School of Psychology, University of New South Wales, Australia

*Corresponding Author
Damian P. Birney
School of Psychology
University of Sydney
Sydney, NSW, 2006, Australia
E-mail: damian.birney@sydney.edu.au

Abstract

The burgeoning increase in the importance given to the role of non-cognitive factors in complex decisions making has led to calls to dethrone intelligence as the primary explanatory model of success. Using a combined experimental-differential paradigm and mixed-level modelling, features of 8 business microworld simulations were experimentally manipulated to investigate the incremental value of non-cognitive predictors of learning and performance trajectories. It was predicted that facilitating personality traits (openness and extraversion), growth/motivational mind sets (learning goals, need for cognition, and beliefs of malleability), and tentatively, emotional-regulation (managing and facilitating emotions) would moderate the impact of complexity and experience on performance and learning trajectories. Results based on a sample of 142 experienced mid-level senior managers suggest microworld simulations can be manipulated to be differentially sensitive to domain-specific differences in reasoning. Reasoning moderated learning effects and the capacity to deal with task complexity. Of the 16 non-cognitive factors investigated, only performance-goal orientations moderated performance trajectories over and above reasoning. These findings give reason to question the importance of non-cognitive dispositional and motivational factors in learning and problem-solving, over and above intellect. We argue that microworld simulations and mixed-level modelling approaches can support the experimental investigations needed for comprehensive, dynamic real-world training research.

Introduction

Success is determined by a multiplicity of factors. Cognitive abilities, such as working-memory and general reasoning capacity, have been demonstrated time and time again as consistent and dominant predictors of work and formal educational outcomes (Gottfredson, 1997; Sternberg & Grigorenko, 2002). Although the role for "non-cognitive" factors in training is rarely disputed, they have historically been considered secondary to intellect. In a review of the work-based training literature of the late 20th century, Colquitt, LePine, and Noe (2000) concluded that traditional cognitive abilities approaches to "… trainability [are] insufficient, given the strong effects of motivational variables over and above cognitive ability" (p. 702). Sitzmann and Ely's (2011) recent meta-analysis of self-regulated learning in work-related training reported that taken together, self-regulatory variables provide significant incremental prediction and largely debunk the ubiquitous status of cognitive abilities as *the* predictor of learning outcomes (17% of the variance in learning was accounted for by goals, persistence, effort, and self-efficacy, after controlling for cognitive ability). As somewhat of a counter-point, Scherbaum, Goldstein, Yusko, Ryan, and Hanges (2012) propose that historical conceptualizations of intelligence remain overvalued and that new conceptualizations are under-researched. They further argue that the importance of intellectual abilities (regardless of their conceptualisation) in supporting learning and success at work is no longer proportionally reflected in the amount of new research into intelligence, relative to other constructs, at least within the I-O literature.

The current research contributes to this debate and extends on previous research that has sought to understand the role of conative dispositions in complex decision-making through investigations of self-regulation (Bandura, 1997; Birney, Beckmann, Beckmann, & Double, 2017; Güss, Burger, & Dörner, 2017; Metcalfe, 1993; Mitchum, Kelley, & Fox, 2016; Stankov, 1999; Stankov & Lee, 2017; Zimmerman, 2002). We achieve this by investigating moderators implicated in self-regulated learning and their association with performance under different experimental manipulations of a microworld simulation.

Based on our review of the literature, it is our contention that under the right conditions, the importance of conative dispositions in the application of cognitive resources to learning, decision making and problem-solving should be observable, if they exist, over and above cognitive abilities. However, the traditional approach to studying correlates of problem solving that almost exclusively relies on between-subjects designs, is limited in its capacity to serve as the basis of an integrated theory of complex decision making. We therefore take up Cronbach's (1957) plea for an integration of experimental methods in the analysis of individual differences (and vice versa) using a business microworld simulation as an experimental tool. In addition, our study goes beyond the typical methodology found in traditional individual differences research by sampling experienced mid-level senior managers. This combination provides an effective methodological paradigm for investigating real-world learning and performance (Beckmann & Goode, 2017; Goode & Beckmann, 2010), while allowing for experimental manipulations that are necessary to isolate the moderating effects of other variables. Thus, rather than modelling performance using only correlational data, our approach is to consider within-subject variability in decision making in the context of training using a repeated-measures design that considers learning and performance processes that co-occur and evolve over time.

In the following sections, we briefly review the nature and rationale of microworld simulations and the cluster of individual differences variables the extant literature would predict to play a role.

Microworld Simulations

Computer-based simulations are frequently used in training when failure is prohibitively expensive or risky in the real world. They allow trainees to explore, make mistakes and learn valuable lessons in virtual environments that are safe and do not impose real costs on the trainee or the organization. The development of a simulation typically starts with an investigation and analysis of the structure of the real-world task with the goal of identifying core processes and procedures, which are then simulated in a virtual environment – the microworld. Microworld simulations (henceforth, *microworlds*) are used in business education to enable students to experiment with different business strategies under a diversity of conditions, and provide a more dynamic and intrinsically engaging training experience than case-study discussions commonly used (Wood, Beckmann, & Birney, 2009). They are often designed to accelerate learning of underlying problem structure by collapsing long periods of history into short periods of simulation time (Funke, 1998).

Wood et al. (2009) have identified a number of key challenges to knowledge acquisition that are common in microworld simulation tasks. They include the (1) knowledge of simulation variables and their values, distributions and semantic context, (2) the relationships between the simulation variables, and (3) an understanding of any noise or randomness that has also been emulated in the model structure to simulate the real environment. Simulation variables can be classified into decision and outcome variables, as well as intervening mediating and moderating variables. *Decision variables* are those for which the problem solver or learner sets the values. *Output variables* form the feedback that is provided. The outputs are the consequences of the input decisions plus effects due to *intervening relationships* within the model. In practice, microworlds are likely to have multiple inputs, including some outside of the direct control of the learner. As in the real-world, both the uncontrolled inputs and *mediator variables* may be unobserved, making it difficult to incorporate them into one's knowledge of the problem structure. Mediators add to complexity if they entail delays between the inputs and the observed outputs. *Moderator variables* further add to the complexity of the structure to be learnt by changing the relationship between decision and outputs variables under different circumstances (Brehmer & Dörner, 1993; Gonzalez, Thomas, & Vanyukov, 2005; Sterman, 2000). For example, decision rules may change as a result of variations in task conditions, such as the differences in actions required when leading a team of skilled and motivated staff versus the actions needed to engage and lead a team of unskilled and unmotivated staff (Goodman, Wood, & Hendrickx, 2004). Lags and dependencies across decision time-points add further complexity.

Moderators of learning and managing complexity

The move toward a more dynamic and comprehensive consideration of a broader range of abilities and conative state and trait dispositions is apparent in both education (Sternberg, 2003) and work-place psychology (Sitzmann & Ely, 2011). Self-regulatory processes havebeen considered important in the development of expertise (Birney, Beckmann, & Wood, 2012; Birney & Sternberg, 2006; Ericsson, 2003) and metacognition, particularly those related to self-beliefs. They have also been seen as critical to achievement (Stankov & Lee, 2017).

Recently, Birney et al. (2017) showed that even in a constrained, standardized task, such as Raven's Advanced Progressive Matrices (APM) test, performance and learning trajectories were reliably associated with a range of non-cognitive factors, over and above general cognitive ability. Two observations from this work are important here. First, being required to self-reflect on the confidence one has in the accuracy of one's response to an APM item, that is to provide confidence ratings, produced negative reactivity relative to a group who did not provide confidence ratings, resulting in significantly poorer performance overall. Second, individual differences in trait neuroticism were positively associated with *performance* trajectories from easy to hard items (after controlling for item-order), but simultaneously negatively associated with *learning* trajectories from the first to the last item (after controlling for item-difficulty). That is, higher Neuroticism tended to facilitate performance as APM items became more cognitively complex, but impede learning from item to item. In a related yet independent study, Double and Birney (2017) showed that under standard APM administration conditions (and after controlling for general cognitive ability), reasoning self-concept did not predict performance differences. However, under conditions where confidence ratings were required, those with higher reasoning self-concept outperformed those in the standard condition, whereas those with lower reasoning self-concept performed more poorly, both in comparison with the standard condition and those with higher self-concept. This suggests task conditions do interact with conative dispositions (in this case self-concept) to impact performance and supports the general call of researchers such as Sitzmann and Ely (2011), for greater consideration of a broader range of conative and facilitating non-cognitive dispositions in complex decision making.

The current work extends on Birney et al. (2017) and Double and Birney (2017) in a number of ways. First, we consider performance in a microworld, a cognitively challenging but more diverse problem-solving task than a standardised intelligence test. Second the microworld activity has an explicitly stated focus on learning, rather than the *implicit* performance framing (Kozlowski & Bell, 2006) that high-stakes cognitive ability assessment tend to have. Further, and unlike the 40mins time limit of the APM, repeated practice on the simulation is encouraged over a period of 1 to 1.5 hrs. Finally, consistent with the expectation that engagement with microworlds draw on a broader array of non-cognitive factors, we consider, in addition to personality, further measures of conative dispositions that have been shown to support learning. These include implicit theories (beliefs about the malleability of intelligence and personality), learning and performance motivations, and emotional regulation.

Consideration of these variables in our battery is based on the general supposition that self-regulation links cognitive abilities, personality and conative dispositions that drive task-focused effort, persistence and metacognition necessary to meet changing situational demands (Bandura, 1997; Beckmann, Wood, & Minbashian, 2010; e.g., Kanfer & Ackerman, 1989; Mischel & Shoda, 1995). We expect the conative variables in particular to moderate the effect of cognitive complexity and learning experience on performance.

Personality

While personality overall has been shown to account for a 9% of total variance in job performance above that which is predicted by cognitive ability (Schmidt & Hunter, 1998), the influence that each of the five personality dimensions (*Openness, Conscientiousness, Extraversion, Agreeableness* and *Neuroticism*) has on job performance is varied. In the Barrick, Mount, and Judge (2001) meta-analysis, only Neuroticism (negatively) and Conscientiousness (positively) were associated specifically with overall job performance. However, various personality factors have been shown to be related to more specific job characteristics. For instance, Agreeableness has been found to predict teamwork performance (Furnham, Jackson, & Miller, 1999); Extraversion has been associated with managerial and teamwork performance (Barrick et al., 2001); and an association between Neuroticism and performance in teams and skilled jobs has been observed (Salgado, 1997). The predictive value of Openness to Experience is somewhat inconclusive in work-place settings (Alessandri & Vecchione, 2012).

While microworlds are developed to simulate real problems, they remain distinct from the world they aim to emulate in important ways. Although more dynamic than traditional reasoning tasks, microworlds are necessarily more structured and often completed individually rather than in teams (though see, Beckmann, Beckmann, Birney, & Wood, 2015, for an investigation of a team-based microworld). In the context of microworld performance, personality associations with measures of intellect more closely associated with learning and knowledge acquisition are relevant. Crystallized intelligence (*Gc*), sometimes referred to by the term “intelligence as knowledge” (Ackerman, 1996), is by definition implicated in learning. Notions of typical intellectual engagement that characterises the volitional, affective, and dynamic nature of real-world cognition (Ackerman, 1994) are also important. To this end, we observe that the only personality factors associated with *Gc* reported in the comprehensive investigation of personality-intelligence associations across the life-span in a sample of over 2000 individuals by Soubelet and Salthouse (2011), was for Openness (positively) and Extraversion (negatively). Openness is of particular interest, given its theoretical link with *Gc* generally (e.g., Ziegler, Danay, Heene, Asendorpf, & Buhner, 2012), with vocational knowledge acquisition specifically (e.g., Ackerman, 1996), and with academic achievement (Stankov, 2013).

Reasoning

The case for the importance of including intelligence as a moderator is clear. In addition to general cognitive ability, we also consider domain-specific reasoning abilities likely to be influential to performance in our specific microworld – verbal reasoning (to understand and evaluate the logic of verbal arguments), numerical reasoning (to make inferences based on numerical data), and abstract reasoning (application of logical rules which govern abstract sequences).

Motivational Mindsets

Motivation to engage in self-regulation during complex problem-solving and decision making is important to the development of flexible as well as routine expertise (Birney et al., 2012). Sitzmann and Ely (2011) identified goals, persistence, effort, and self-efficacy as the most influential self-regulatory variables in learning. Personal goals and self-efficacy are also well-established predictors of learning and performance in complex tasks (Wood & Bandura, 1989; Wood, Whelan, Sojo, & Wong, 2012). *Goal orientation* research most commonly addresses the distinction between mastery (learning) and performance orientations (Pintrich, 2000; VandeWalle, 1997). In general, a mastery orientation, which refers to a preference for goals that focus on developing skills and improving one's understanding, has been positively linked with self-regulation, and consequently performance (e.g. Middleton & Midgley, 1997). Conversely, a performance orientation, which refers to an approach that focuses on increasing performance against external criteria (e.g., others), has been found to have a negative impact on performance (e.g. Bouffard, Boisvert, Vezeau, & Larouche, 1995; Dweck & Leggett, 1988; Heslin, Latham, & VandeWalle, 2005). This is thought to happen because limited attentional resources are devoted to thoughts unlikely to facilitate self-regulated learning, such as ruminating about what others might think of their performance (Steele-Johnson, Beauregard, Hoover, & Schmidt, 2000). Performance orientations have been further fractionated into two aspects. *Performance-prove* orientated individuals are motivated to obtain favourable performance judgments – to be well evaluated; whereas individuals with a *performance-avoid* orientation tend to avoid potentially unfavourable performance judgments (Heslin et al., 2005; VandeWalle, 1997).

Dweck (2000) has argued that the *implicit theory* that people have about human attributes (whether they perceive highly valued personal attributes such as intelligence or personality, as malleable or not), structures the way they understand and react to human actions and outcomes (both their own and others), and are thus closely linked to goal orientations but more grounded in personal beliefs and attitudes.

Another related construct is *Need for Cognition*, which refers to individual differences in a disposition to want to understand and structure the world (Cacioppo & Petty, 1982), and a tendency to engage in and enjoy effortful cognitive activities. Need for cognition accounts for significant variance in learning over and above cognitive ability (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Consistent with Dweck (2000), we see Need for Cognition, learning/mastery goal orientations and incremental implicit theories to support a *growth* mind set facilitative of performance and learning within a dynamic task.

Emotional Regulation

Theoretical developments in emotional intelligence research (e.g., Mayer & Salovey, 1993) have led to an 'abilities' model composed of four branches: (1) Emotion perception, the ability to accurately perceive and express emotions; (2) Emotional facilitation, the ability to use emotions to facilitate thought and to generate particular emotions when required; (3) Emotion understanding, the ability to interpret meanings of emotions and circumstances, and to understand emotional transitions; and (4) Emotion management, the ability to manage emotions in the self and others, and to recognise the actions that are most effective in improving negative moods and maintaining positive ones. Within the I-O literature, emotional intelligence has been touted as providing advantage in leadership situations and life generally via effective regulation of the emotions and behaviours of oneself and others (Harms & Credé, 2010). While links with learning and performance in cognitive tasks are more circumspect (Brody, 2004), we consider emotional intelligence for two reasons. First as an anchor to the basic and applied I-O research, and second with the view that if EI is to have a process link with cognitive performance, it is most reasonably to occur through the emotional management and facilitation facets.

Aims and Hypotheses

The focus of our investigation is on the profile of individual differences that moderate performance and learning trajectories as a function of experimental manipulations of microworld complexity. Following Birney et al. (2017), the moderation of the influence of task complexity on performance by a personality-related individual differences variable (e.g., openness) is referred to as a *psychometric complexity* (ψ_C) effect. A variable which moderates performance as a function of experience is referred to as having a *psychometric learning* (ψ_L) effect. Our guiding expectation is that conative variables are the most likely candidates for psychometric complexity and learning moderators of performance. The main research hypothesis is that facilitating personality traits (Openness), growth/motivational mind sets (learning goals, need for cognition, and malleability beliefs of ability), and emotional-regulation (managing and facilitating emotions) will account for incremental variance in individual microworld performance and learning trajectories.

Method

Participants

In total, 162 mid-level managers participating in a long-term training and development program designed to improve leadership skills, took part in this study. The final sample, with complete individual differences data (N = 142, 88 Male), consisted of managers working across a variety of organisations including a bank (N = 41), an international airline (n = 36) and an insurance company (n = 63), with a mean age of 33.14 years (SD = 5.15) and mean years of management experience of 4.43 (SD = 3.59). In comparison to a normative sample of managers, participants were typical in terms of domain-specific verbal and numerical reasoning abilities (; t_{140} = .91, p = .36, and t_{140} = .50, p = .62, respectively), but higher than typical in abstract reasoning (t_{140} = 6.93, p < .001). Variability in each reasoning domain was typical of the normative sample and comparable across domains.

Materials

Experimental Task: Inventory and Workforce Management Simulation

The microworld modelled standard business stock management processes and decision making. The theoretical complexity of decisions was manipulated along two independent dimensions intrinsic to the task, *delays* and *outflow*. Delays occurred with regard to hiring and firing decisions (due to time needed to train new hires or due to notice periods when firing). Outflow of stock, over and above sales (e.g., through waste, defects, etc), was another complexity relevant variable. The goal for the participant was to reach and maintain an ideal level of net inventory by taking into consideration staffing delays and stock outflow over a simulated period of 30 weeks via the management of the workforce (number of staff). The net inventory is thus managed solely via weekly staff hiring and firing decisions. Each weekly decision constitutes a ‘trial’ within the microworld. A ‘run’, consisting of a maximum of 30 trials, constitutes the 30-week simulated period. Performance was conceptualised as a penalty score associated with the costs of a suboptimal level of inventory and staffing. A graph of the trial-by-trial, accumulating penalty score is refreshed after each trial and serves as one source of feedback to the participant (see Figure A1 in Appendix A).

Experimental manipulation of complexity: Delays have a knowable relational structure – a greater lag between decisions and their impact generates a concomitant increase in cognitive demand, which greater information processing capacity (in form of reasoning ability) is expected to mitigate. Variable (i.e., random around a mean) *outflow*, on the other hand, results in unpredictable deviations from a targeted stock level. Reasoning ability, in this instance, is expected to be less effective in mitigating the impact of randomness on performance. In short, since the source of difficulty of delays and outflow are different, we expect reasoning to moderate the impact of delays, but not outflow.

We created 8 variants of the simulation that combined one of four levels of delay complexity (0, 1, 3, and 5 weeks) and one of three levels of outflow complexity (none, constant at 10 units per week, and varied with an average of 12 units per week). Given consideration of training time and fatigue, it was not possible to completely cross all levels of delay with all levels of outflow. The variants used are summarised in Table 1.

Analysis strategy: Given the incomplete design, the separate delay and outflow complexity levels were collapsed into control vs load dichotomies for analyses. For delay, the *load* condition comprised the 1, 3 and 5 week levels (SM02, SM03, SM04, SM06, SM08), and the *control* condition was ‘no delay’ (SM01, SM05, SM07). For outflow, the *load* condition comprised the varied levels (SM07, SM08) and the *control* condition comprised the ‘none’ and ‘constant’ outflow levels (SM01-SM06). Psychometric complexity would be observed in statistically significant cross-level interactions between delay and outflow, and moderating variables on performance, respectively. We define learning *experientially* as the number of simulation attempts, and thus psychometric learning would be observed in statistically significant cross-level interactions between number of attempts and the moderators. More detail on our analysis strategy is provided in the Results section.

[Insert Table 1 about here]

Microworld administration: The focus of the microworld activity was on learning about delays and non-linearities associated with constant versus varied outflow. The general principle we adopted for administration is standard in education, and was as follows: At the beginning of the session, the facilitator explained the rationale for the activity and provided participants with an information sheet summarising the explanation (see Appendix A). Participants were then presented with the most difficult microworld (SM08) to orient them to the ultimate challenge (where the combination of delays and outflow complexity was greatest), and to contextualise the learning rationale of the simplified variants. Participants were encouraged to work through the remaining microworlds in the order they were listed on the computer display (SM01 to SM08) with the objective to move toward repeated attempts of SM08 by the end of the 1.5 hr session. In practice, it was the participants’ choice as to how many runs of each microworld were attempted. Thus, all participants attempted SM08 first and then worked through a variable number of runs of SM01 to SM08 in an order of their choosing.

Personality

Goldberg's 50-item self-report version of the International Personality Item Pool (IPIP, http://ipip.ori.org/) was used to assess the Big Five personality factors: Openness to Experience ($\alpha = .78$), Extraversion ($\alpha = .88$), Agreeableness ($\alpha = .76$), Conscientiousness ($\alpha = .87$) and Neuroticism ($\alpha = .85$). Each factor was assessed with 10 items which described personal characteristics (e.g. "am always prepared"). Participants rated how accurately the items generally described them by using an anchored scale from "very inaccurate" (scored as 0) to "very accurate" (scored as 100).

Reasoning Ability

Verbal Reasoning – VR (SHL – VMG4) is a 48-item commercially-sourced test (shl.com) that measures the ability to understand and evaluate the logic of various verbal arguments relevant to managerial work. The task was to decide whether a statement made in connection with the given information was true, untrue, or whether there was insufficient information to judge (Cronbach $\alpha = .82$).

Numerical Reasoning – NR (SHL – NMG4) is a 35-item commercially-sourced test (shl.com) measuring the ability to make decisions, or inferences, based on numerical data and was designed to apply to a range of management level jobs. The task was to interpret data and combine information from different sources in order to answer the questions given. Calculators were provided to ensure emphasis on understanding and evaluating rather than on computation (Cronbach $\alpha = .91$).

Abstract Reasoning – AR (SHL – DC3.1) is a 40-item commercially sourced test (shl.com) measuring the ability to reason with abstract figures. It requires the recognition and application of logical rules which govern sequence changes. The abstract reasoning test consisted of a series of diagrammatic sequences whereby participants are required to identify the underlying structure of sequence and select the figure that best completed the pattern (Cronbach $\alpha = .85$).

Reasoning Composite – Principal components analysis of the reasoning measures was used to derive a single measure. The first principal component captured 63.25% of the common variance in the reasoning subtest scores. Component loadings were *VR* = .743, *NR* = .872, and *AR* = .766.

Motivation

Goal orientation refers to the framing of tasks in terms of opportunities to either learn about the task or to perform. Three goal orientations were assessed using the VandeWalle (1997) instrument. *Learning Orientation (LGO)* was assessed with six items (α = .82); *Performance-Prove Orientation (PGP)* was assessed with five items (α = .56); and *Performance-Avoid Orientation (PGA)* was assessed with five items (α = .78).

Need for Cognition (NC) was assessed using the 18-item version (α = .83) of the Cacioppo and Petty (1982) Need for Cognition Scale. The scale refers to one's tendency to engage in and enjoy effortful cognitive endeavours.

Implicit Theories - Intelligence (ITI) and Personality (ITP) refer to beliefs individuals hold about the changeability of intelligence. The *Implicit Theories Scale* which contained 4 items for intelligence (α = .90) and 4 items for personality (α = .83) were taken from Dweck (2000). High scores reflect incremental (changeability) beliefs and low scores reflect entity (fixedness) beliefs.

Emotional Intelligence

Participants completed the *Mayer-Salovey-Caruso Emotional Intelligence Test* (MSCEIT). Standardized specific task scores for 1) *Perceiving Emotions*, 2) *Using Emotions (Facilitating thought)*, 3) *Understanding Emotions*, and 4) *Managing Emotions* were derived using consensus-group of managers as the criterion norms.

Procedure

The microworld simulation and individual assessments were presented to participants in cohorts of between 10-15 during Module 1 of the Accelerated Learning Laboratory leadership development program (see Appendix C). Over the course of three days, participants completed the conative individual differences measures, then participated in the microworld simulation activities and then completed the cognitive ability measures.

Results and Discussion

Overview of Analyses

We analyse the data across three nested (i.e., incremental) models. The lowest unit of analysis is the run. Attempt number was derived on the total data and then runs for which there were 5 or fewer trials excluded as non-serious attempts[1]. Model 1 tests a baseline model of the overall effect of the run-based *complexity* factors (delay and outflow) and simultaneously the run-based person-specific *experience* factor (number of attempts) on performance. We expect greater delay and variable outflow to be associated with higher penalty scores (i.e. lower performance), and increased number of attempts associated with lower penalty scores (i.e., higher performance, learning). Model 2 includes the *Reasoning* composite as a cross-level predictor of mean performance (i.e., modelling varying intercepts). We expect higher general reasoning ability to be associated with lower penalty scores. Model 3 additionally includes cross-level interactions between the complexity (delay and outflow) and experience (number of attempts) factors and the *moderators* (i.e., modelling slopes). The expectation is that the reasoning and conative variables will moderate the complexity and experience effects – specifically, to be associated with a buffering psychometric complexity effect, and to facilitate a psychometric learning effect (as a function of experience), as defined earlier. To investigate moderation of domain-specific reasoning, Model 3* replaces

[1] Not all participants completed all 30 trials on all runs. One analytic strategy would be to include only complete runs for analysis. However, this would obscure the learning that would occur on the incomplete attempts. Instead, runs with more than 5 trials were included and the standardised number of trials completed was entered as a level 1 covariate in all analyses.

the reasoning composite with the domain-specific reasoning variables and includes each as a moderator in three separate additional analyses.

All analyses were conducted using *R* version 3.3.2 and Linear Mixed Effects (LME) modelling was performed using the *lme4* (Bates, Maechler, Bolker, & Walker, 2017) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016) packages. We report standardized regression coefficients and 95% confidence intervals to assist with evaluation of effect sizes. Detailed results are reported in Appendix B.

Descriptive Statistics

In terms of anchoring our sample of managers with other studies,
Table 2 reports the run-based descriptive statistics for the eight variants of the inventory management microworld. In total, there were 2116 unique runs across 142 participants available for analysis. Participants on average completed 17.5 (SD = 6.57) runs. The number of unique participants who completed a given microworld ranged from 92 to 140.

[Insert Table 2 about here]

Moderators of performance: Complexity and learning

A linear mixed-effects model (*Model 1: Baseline*) was tested by regressing the run-specific penalty score as the dependent variable on the main effects and interaction between outflow and delay, controlling for number of trials per run and run attempt number. More trials per run resulted in significantly lower penalty scores, $\beta = -0.25$, *CI* [-0.29 | -0.22)], $p < .001$. There is evidence of learning overall as a function of experience – increasing number of attempts (controlling for delay and outflow difficulty, and trials attempted) was associated with lower penalty scores, $\beta = -0.24$, *CI* [-0.28 | -0.19], $p < .001$. As expected, delays and variable outflow were each associated with significantly higher penalty scores, $\beta = 0.26$, *CI* [0.19 | 0.33], $p < .001$ and $\beta = 0.17$, *CI* [0.09 | 0.24], $p < .001$, respectively). These task main-effects were qualified by a significant interaction. As expected, the effect of delay was more pronounced when outflow was variable, β = 0.18, *CI* [0.09 | 0.26], $p < .001$. The intra-class correlation (ICC = .168) suggests that 16.8 % of the total variability in penalty score is due to differences between individuals, and 83.2% due to differences between runs. Model 1 accounted for approximately 43% of the variability in penalty scores across levels ($\Omega^2 = .435$) [2].

The second analysis (*Model 2: Baseline + Reasoning covariate*) extended on Model 1 by including the reasoning composite as a covariate on the intercept (the individuals' mean penalty score). The significant effects of Model 1 were maintained, and *Reasoning* was a significant incremental predictor such that higher reasoning ability was associated with lower penalty scores overall (β = -0.19, *CI* [-0.24 | -0.13], $p < .001$). Model 2 accounted for approximately 51% of the variability across the fixed and varying components of the penalty score ($\Omega^2 = .510$), and according to a test on the change in deviance ($\chi^2_{(10)} = 104.59$, $p < .001$), was a significant improvement on Model 1.

In order to investigate the incremental value of the individual differences variables, Model 3 (*Baseline + Reasoning Composite + Moderator*) considered each moderator in cross-level interactions with task complexity and number of attempts (learning opportunity), while controlling for the influence of *Reasoning* on overall performance (a separate analysis for each moderator). As a comparison, Table 3 reports the simple correlation between the individual differences variables and the person-level intercept and slope estimates extracted from Model 1.

[Table 3 about here]

Reasoning significantly moderated the delay effect, β = -0.19, *CI* [-0.24 | -0.13], $p < .001$, such that higher Reasoning performance was associated with a less pronounced cost of

[2] Traditional estimates of R^2 have problematic interpretations due to the cross-level interaction of fixed and random effects. In response, a variety of pseudo R^2 estimates have been proposed. The estimate reported here is as described by Nakagawa and Schielzeth (2013) and referred to there as omega-squared (Ω^2). It is implemented in the R package *sjstats* (Lüdecke, 2017).

a delay in decision making. This cross-level interaction is represented in Figure 1A. The outflow effect and experience (number of attempts) was not moderated by *Reasoning*. Model 3 (*baseline + Reasoning covariate + Reasoning moderator*) was a statistically significant improvement on Model 2, $\chi^2_{(3)} = 10.72$, $p = .013$.

The analyses of the personality, motivation, and the emotional regulation variables indicated that they were, by and large, unable to account for any further variation in performance, over and above general reasoning ability. The only caveat to this was the significant main-effects and psychometric complexity effects for the performance goal orientation variables. Higher *performance-avoid* orientation was associated with an overall lower penalty score (i.e., better performance), $\beta = -0.14$, *CI* [-0.25 | -0.03], $p = .014$. This effect was however qualified by a significant interaction with outflow, such that higher dispositions toward demonstrating performance to avoid negative evaluations, was associated with an increasingly pronounced performance benefit in the control condition (0 and constant outflow), but not in the variable outflow load condition (see Figure 1B). This is a reverse ψ_C (psychometric complexity) effect, because higher scores were associated with better performance in low complexity items, whereas ψ_C is formalised as buffering moderation (Birney et al., 2017). There was similarly a significant but reversed ψ_C effect of *performance-prove* goal orientations on outflow, $\beta = 0.11$, CI [0.02 | 0.20], $p = .024$. Again, higher performance-prove orientation was associated with a pronounced benefit on the control outflow levels (Figure 1C), suggesting a relatively more pronounced focus on microworlds with easier outflow levels where performance might more readily be demonstrated.

As noted, ψ_C effects of both performance goals are not facilitating ones. Unlike the cognitive ability ψ_C effect, which was associated with a benefit in meeting the challenge of the more difficult task manipulations (relative to the control condition) with increased *Reasoning* ability, performance on the more difficult, varied outflow condition was not facilitated by performance goal orientations. This is, however, theoretically consistent with the extant literature on mindsets (Dweck, 2000; Heslin et al., 2005; VandeWalle, 1997), where it is suggested that performance goal orientation is associated with a focus on aspects of tasks where capability can be demonstrated and failure avoided – in our case, the easiest levels (control) of the easiest complexity manipulation (outflow).

There were also some theoretically interesting ψ_C (delay) trends for another motivation variable, *Need for Cognition*, however these did not reach statistical significance ($p \approx .06$). We note this because it was the only other moderator to approach an effect-size comparable to the reasoning variables. *Need for Cognition* (Cacioppo et al., 1996) has been associated with the typical intellectual engagement (Ackerman, 1994) that is more representative of work-based learning from experience and microworld engagement (Güss et al., 2017). Taken together, further investigation seems warranted. No statistically significant effects were observed for any of the personality variables nor for the emotional regulation variables. Also, contrary to expectations, no moderation of learning (cross-level interaction with number of attempts) was observed.

Model 3* analyses were conducted to investigate psychometric complexity and learning effects associated with the domain-specific reasoning variables (*Abstract*, *Verbal,* and *Numeric*), in comparison to the domain-general composite. As would be expected, the results are in accordance with analyses of the composite *Reasoning* measure (e.g., higher scores in the specific domains were associated with a less pronounced costs of delays). However, there were two important departures from the analysis of the domain-general composite. First, while *Verbal* reasoning ($\beta = -0.14$, *CI* [-0.26 – -0.03], $p = .017$) and *Numerical* reasoning ($\beta = -0.18$, *CI* [-0.29 – -0.06], $p = .003$) were significant predictors of lower overall penalty scores, *Abstract* reasoning was not ($\beta = -0.09$, *CI* [-0.20 – 0.03], $p = .131$). Second, *Abstract* reasoning did not moderate learning trajectories (cross-level interaction with number of attempts), $\beta \approx 0.00$, *CI* [-0.07 – 0.06], $p = .919$, however this was not the case for the other specific reasoning domains. Higher *Verbal* reasoning was associated with a significantly more pronounced benefit of experience – a psychometric learning effect – as realised in lower penalty scores with increasing number of attempts, $\beta = -0.07$, *CI* [-0.13 – -0.00], $p = .047$. A similar effect for *Numerical* reasoning failed to reach statistical significance, $\beta = -0.06$, *CI* [-0.12 – 0.01], $p = .076$). The interactions are plotted in Figure 2. These results tentatively suggest that domain-specific verbal ability is particularly important in translating microworld learning opportunities into performance gains. When *Abstract* and *Numeric* reasoning are included as main-effect covariates in the analysis with *Verbal* reasoning, the Verbal reasoning effects remain uniquely significant (delay moderation, ψ_C: $\beta = -0.07$, *CI* [-0.15 – -0.00], $p = .038$, Attempt moderation, ψ_L: $\beta = -0.07$, *CI* [-0.13 – -0.01], $p = .036$) over and above the other domain-specific reasoning. This was not the case in the comparable analyses of Numeric reasoning (controlling for Abstract and Verbal), or for Abstract reasoning (controlling for Numerical and Verbal), suggesting an important role for domain-specific verbal reasoning in this microworld.

[Figure 2 about here]

General Discussion

We began our investigation based on the contention that under the right conditions, the importance of conative dispositions in the application of cognitive resources to learning and problem-solving should be observable, over and above general cognitive ability (Birney et al., 2017; Colquitt et al., 2000; Double & Birney, 2017; Sitzmann & Ely, 2011; Wood et al., 2009). To this end we employed a combined experimental and differential approach to study a broad array of factors thought to be variously related to learning trajectories, to investigate performance in a simulation of a real-world stock management scenario. Our microworld activity was explicitly framed with learning goals (Kozlowski & Bell, 2006) and implemented within a structured training program. Rate of progression and the extent of practice was guided by program facilitators, but ultimately determined by the trainees themselves. Trainees were current middle-level managers from multiple industries who had been nominated by their company to participate in a leadership development program. The study environment thus represents a personally relevant real-world learning context, for "real" participants who have real intentions to learn. The use of a microworld served multiple purposes, but primarily it was chosen to provide a sufficiently dynamic but necessarily structured framework for investigation of engaged learning and performance under experimentally controlled complexity conditions.

Our results are clear, albeit somewhat sobering for the "non-cognitive" case. In the remaining sections, we reflect on the implications of these findings to the field, and suggest avenues for future investigations.

Implications

Reasoning: Domain-general versus domain-specific

Consistent with expectations, reasoning was a dominant predictor of performance. Higher reasoning scores in all three specific domains, as well as the domain-general score, was associated with better performance in meeting the challenge of the increased complexity of the control versus delay conditions. This suggests that the complexity of the delay manipulation is (in part) due to an increased demand on reasoning abilities, consistent with the presence of psychometric complexity (Birney et al., 2017). On the other hand, there was no psychometric complexity effect for outflow. Higher reasoning ability did not afford a differential benefit in managing the complexity of constant (i.e., predictable) versus variable (i.e, unpredictable) outflow; or phrased differently, the outflow complexity manipulation had an impact on difficulty that could not be explained by individual differences in reasoning ability. Hence our prediction that difficulty caused by delays or outflow rates tap different constructs is supported. Delays appear to be primarily cognition related, whereas the unpredictability of variable outflow rates may be related to non-cognitive factors.

We also observed different effects as a function of the type of domain-specific reasoning considered. Relative to *Abstract* and *Numerical* reasoning, *Verbal* reasoning ability in particular showed pronounced overall effects, psychometric complexity effects, as well as psychometric learning effects. Taken at face-value, these differences suggest our microworld is sensitive to different domain-specific reasoning demand.

Conative dispositions and traits

Goals to perform so as to demonstrate competence or avoid negative evaluations were associated with an increasingly pronounced performance benefit in the less complex control outflow condition, but not in the variable condition. Similar results were found for those with performance-prove orientation. This is theoretically consistent with the predicted interpretation that performance goal orientations are associated with a focus on aspects of tasks where capability can be demonstrated and failure avoided (Dweck, 2000; Heslin et al., 2005; VandeWalle, 1997). Alternatively, a focus on task components that are knowable (and lead to predictable decision-outcomes), rather than not (as in the case of varied outflow) may be advantageous in terms of resource allocation, and a positive outcome of a performance goal orientation. Replication of these findings, along with the differentiation of reasoning domain-specificity, would provide further evidence of the utility of microworlds as a paradigm for investigating cognitive and non-cognitive processes in the same task.

Contrary to expectations, no comparable effects were associated with what we referred to as conative dispositions toward learning. Facilitating personality traits (openness), growth/motivational mind sets (learning goals, and incremental beliefs of ability), and emotional regulation (managing and facilitating emotions) did not account for incremental variance in individual performance and learning trajectories.

Limitations

It was surprising to us that the non-cognitive measures were so unimpressive in predicting learning and performance in a dynamic learning / performance situation. Theoretical justifications for observing an effect over and above reasoning abilities are reasonable (Ackerman, 1994; Birney et al., 2012; Colquitt et al., 2000; Sitzmann & Ely, 2011). Thus, while there is promise in the use of microworld simulations, we may not yet have the right conditions in place to observe the impact of non-cognitive processes on performance. There are a number of reasons why we may have failed to find significant moderation effects, including a potential lack of power, (un)reliability of measurement, and noise added by insufficient structure in the training guidelines, which we consider in turn. First, power to detect effects for cross-level interactions with slopes (effectively a 3-way interaction) is typically low relative to main-effects, and thus a meaningful concern. However, the risk of insufficient statistical power would be greater if there was some evidence that the simple bivariate relationships between microworld complexity and learning effects and the conative variables approached some meaningful value (Stankov & Lee, 2014), but they did not. Thus, we are relatively confident that an increase in statistical power will not change the result pattern dramatically.

Second, the individual differences variables have been measured via published, validated tests, and their reliability in the current sample is good. Microworld simulations on the other hand, are complex and dynamic, and the challenge of extracting an appropriate (valid) and reliable performance metric is not a trivial one (Wood et al., 2009). In the current work, we demonstrate evidence of systematicity in the performance metric used – approximately 43% of the run-level variability in Penalty score is accounted for by run-level task (outflow and delay) and person (number of trials and runs attempted) factors. However, this still leaves a substantial amount of variance in performance scores unaccounted for. We have modelled four different metrics of performance, the mean penalty (as varying intercepts), as well as the slopes of delay, outflow, and number of attempts. While this provides a rich set of metrics with different theoretical implications, their effectiveness remains linked to the appropriateness of the run-based penalty score.

The third issue raised was that the training environment itself introduced noise to the experiment which has obscured identification of true effects – this is a further issue related to power. The point here is the tension between testing efficacy (as a form of proof of concept under ideal conditions) and testing effectiveness (as a test of effects under pragmatic or “realistic” conditions). While it is reasonable to be concerned about the rigor of the experiment, our methods must be flexible enough to deal with a real-world training environment if we are to get a handle on the actual moderators of motivated real-world adult learning and performance. This remains a challenge for the field (Güss et al., 2017).

Were to from here?

An alternative explanation is that the importance of non-cognitive dispositional and motivational factors in learning and problem-solving, over and above cognition, has been oversold. Our findings certainly give reason for pause, at least within the context of a simulated real-world business scenario. Our investigations of microworld simulations suggest they have methodological potential to balance real-world fidelity, scientific rigour, and an engaging learning context (Beckmann et al., 2015), and may be experimentally manipulated to be sensitive to cognitive as well as non-cognitive drivers of performance. Rather than discount non-cognitive variables, our findings speak to the importance of intellect and support our ongoing search for the right conditions to observe their influence.

References

Ackerman, P. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence: Theories of Intelligence* (Vol. 4).

Ackerman, P. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22*, 227-257.

Alessandri, G., & Vecchione, M. (2012). The higher-order factors of the Big Five as predictors of job performance. *Personality and Individual Differences, 53*, 779+784.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Barrick, M., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9-30.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2017). lme4: Linear Mixed-Effects Models using 'Eigen' and S4. https://cran.r-project.org/web/packages/lme4/lme4.pdf. .

Beckmann, J. F., & Goode, N. (2017). Missing the wood for the wrong trees: On the difficulty of defining the complexity of complex problem solving scenarios. *Journal of Intelligence, 5*, 15.

Beckmann, N., Beckmann, J. F., Birney, D. P., & Wood, R. E. (2015). A problem shared is learning doubled: Deliberative processing in dyads improves learning in complex dynamic decision-making tasks. *Computers in Human Behavior, 48*, 654-662.

Beckmann, N., Wood, R. E., & Minbashian, A. (2010). It depends how you look at it: On the relationship between neuroticism and conscientiousness at the within- and the between-person levels of analysis. *Journal of Research in Personality, 44*, 593-601.

Birney, D. P., Beckmann, J. F., Beckmann, N., & Double, K. S. (2017). Beyond the intellect: Complexity and learning trajectories in Raven's Progressive Matrices depend on self-regulatory processes and conative dispositions. *Intelligence, 61*, 63-77.

Birney, D. P., Beckmann, J. F., & Wood, R. E. (2012). Precursors to the development of flexible expertise: Metacognitive self-evaluations as antecedences and consequences in adult learning. *Learning and Individual Differences, 22*(5), 563-574.

Birney, D. P., & Sternberg, R. J. (2006). Intelligence and cognitive abilities as competencies in development. In E. Bialystok & G. Craik (Eds.), *Lifespan Cognition: Mechanisms of Change* (pp. 315-330). New York: Oxford University Press.

Bouffard, T., Boisvert, J., Vezeau, C., & Larouche, C. (1995). The impact of goal orientation on self-regulation and performance among college students. *British Journal of Educational Psychology, 65*(3), 317-329.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior, 9*, 171-184.

Brody, N. (2004). What cogntive intelligence is and what emotional intelligence is not. *Psychological Inquiry, 15*(3), 234-238.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116-131.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197-253.

Colquitt, J. A., LePine, A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*(5), 678-707.

Double, K. S., & Birney, D. P. (2017). Are you sure about that? Eliciting metareasoning judgements improves performance on Raven's progressive matrices. *Thinking & Reasoning, 23*(2), 190-206.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive apporach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Ericsson, K. A. (2003). The acquisition of expert performance as problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 31-83). Cambridge, UK: Cambridge University Press.

Funke, J. (1998). Computer-based testing and training with scenarios from complex problem-solving reseach: Advantages and disadvantages. *International Journal of Selection and Assessment, 6*(2), 90-96.

Furnham, A., Jackson, C. J., & Miller, T. (1999). Personality, learning style and work performance. *Personality and Individual Differences, 27*(1113-1122).

Gonzalez, C., Thomas, R. P., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence, 33*, 169-186.

Goode, N., & Beckmann, J. F. (2010). You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence, 38*, 345-352.

Goodman, J. S., Wood, R. E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology, 89*, 248.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence, 24*(1), 79-732.

Güss, C. D., Burger, M. L., & Dörner, D. (2017). The role of motivation in complex problem solving. *Frontiers in Psychology, 8*(851), 1-5.

Harms, P. D., & Credé, M. (2010). Emotional intelligence and transformational and transactional leadership: A meta-analysis. *Journal of Leadership & Organizational Studies, 17*(1), 5-17.

Heslin, P. A., Latham, G. P., & VandeWalle, D. (2005). The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology, 90*, 842-856.

Kanfer, R., & Ackerman, P. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(2), 657-690.

Kozlowski, S. W. J., & Bell, B. S. (2006). Disentangling achievement orientation and goal setting: Effects on self-regulatory processes. *Journal of Applied Psychology, 91*(4), 900-916.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in Linear Mixed Effects Models. https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf.

Lüdecke, D. (2017). Collection of convenient functions for common statistical computations. https://cran.r-project.org/web/packages/sjstats/sjstats.pdf.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence, 22*(89-114).

Metcalfe, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff Amnesia. *Psychological Review, 100*, 3-22.

Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of educational psychology, 89*(4), 710.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2), 246-268.

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General, 145*(2), 200-219.

Nakagawa, A., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142. doi: doi: 10.1111/j.2041-210x.2012.00261.x

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning.

Salgado, J. F. (1997). The five factor model of personality and Job Performance in the European Community. *Journal of Applied Psychology, 82*(30-43).

Scherbaum, C. A., Goldstein, H. W., Yusko, K. P., Ryan, R., & Hanges, P. J. (2012). Intelligence 2.0: Reestablishing a research program on g in I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 5*(2), 128-148.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421-442. doi: 10.1037/a0022777

Soubelet, A., & Salthouse, T. A. (2011). Personality-cognition relations across adulthood. *Developmental Psychology, 47*(2), 303-310.

Stankov, L. (1999). Mining on the "No Man's Land" between intelligence and personality. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 315-337). Washington, DC: American Psychological Association.

Stankov, L. (2013). Noncognitive predictors of intelligence and academic achievement: An important role of confidence. *Personality and Individual Differences, 55*, 727-732.

Stankov, L., & Lee, J. (2014). Quest for the best non-cognitive predictor of academic achievement. *Educational Psychology, 34*(1), 1-8.

Stankov, L., & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence, 61*, 11-16.

Steele-Johnson, D., Beauregard, R. S., Hoover, P. B., & Schmidt, A. M. (2000). Goal orientation and task demand effects on motivation, affect, and performance. *Journal of Applied Psychology, 85*(5), 724-738.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world* London, USA: McGraw-Hill

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2002). *The general factor of intelligence: How general is it?* Mahwah, NJ: Lawrence Erlbaum Associates.

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement, 57*, 995-1015.

Wood, R. E., & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology, 56*, 407-415.

Wood, R. E., Beckmann, J. F., & Birney, D. P. (2009). Simulations, learning and real world capabilities. *Education + Training, 51*, 491-510.

Wood, R. E., Whelan, J., Sojo, V., & Wong, M. (2012). Goals, goal orientations, strategies and performance. In E. A. Locke & G. P. Latham (Eds.), *New Developments in Goal Setting and Task Performance.* San Francisco: Routledge.

Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Buhner, J. (2012). Openness, fluid intelligence, and crystallized intelligence: Toward an integrative model. *Journal of Research in Personality, 46*, 173-183.

Zimmerman, B. J. (2002). Achieving academic excellence: A self-regulatory perspective. In M. Ferrari (Ed.), *The pursuit of excellence through education* (pp. 85-110). Mahwah, NJ: Lawrence Erlbaum Associates.

Table 1.

Experimental design of the experimental conditions of the microworld simulation

Complexity Manipulations*		**Outflow** None (0)	Constant (0)	Varied (1)
No delay	(0)	SM01	SM05	SM07
1 week delay	(1)	SM02	-	-
3 weeks delay	(1)	SM03	SM06	SM08
5 weeks delay	(1)	SM04	-	-

* Values in parentheses represent the complexity coding of the control vs experimental condition used in analyses.

Table 2.

Descriptive statistics. Mean penalty score (SE in parentheses) by simulation across all data points

Microworld	**Outflow**		**Delay**		**Runs**	**Penalty**	**SE**	**Unique Participants**	**Mean Attempts**
SM01	0	{0}	0	{0}	126	122.29	(27.70)	92	1.37
SM02	0	{0}	1	{1}	248	631.98	(45.54)	138	1.80
SM03	0	{0}	3	{1}	176	554.56	(48.17)	123	1.43
SM04	0	{0}	5	{1}	133	492.91	(44.35)	108	1.23
SM05	C	{0}	0	{0}	126	118.77	(31.60)	96	1.31
SM06	C	{0}	3	{1}	428	819.62	(30.69)	138	3.10
SM07	V	{1}	0	{0}	278	257.39	(21.26)	139	2.00
SM08	V	{1}	3	{1}	601	884.00	(25.17)	140	4.29
Total number of observations =					**2116**	**Total Participants =**		**142**	

Notes: Outflow 0 = no outflow, C = constant at 10 units per week; V = varied with 12 units per week on average; Delay = in weeks; Runs = number of unique runs in the data; values in brackets represent the coding for analyses. Participants = unique number of participants who completed the simulation; Mean Attempts = average number of attempts per participant.

Table 3.

Zero-order correlations between Model 1 person-level parameters (varying intercepts and varying slopes) of microworld performance and individual differences variables

Moderator	Mean Penality (Intercept)	Attempts (learning slope)	Outflow (complexity slope)	Delay (complexity slope)
Verbal Reasoning	-0.317 **	-0.200 *	0.021	-0.358 **
Numerical Reasoning	-0.361 **	-0.199 *	0.000	-0.446 **
Abstract Reasoning	-0.232 **	-0.048	0.013	-0.293 **
Reasoning Composite	-0.381 **	-0.188 *	0.013	-0.461 **
Openness	0.041	0.076	-0.048	-0.023
Conscientiousness	0.137	0.102	-0.097	0.043
Agreeableness	0.092	0.026	0.015	0.140
Extraversion	0.003	0.056	-0.020	-0.037
Neuroticism	0.006	0.024	0.046	0.057
Learning Goals	0.133	0.006	-0.111	0.049
Performance-Prove Goals	-0.130	-0.125	0.170 *	0.060
Performance-Avoid Goals	-0.244 **	-0.094	0.224 **	-0.043
Implicit Theories - Intelligence	0.112	0.060	0.042	0.189 *
Implicit Theories - Personality	0.020	0.020	0.005	0.029
Nead for Cogntion	-0.016	-0.017	-0.064	-0.095
Emotional Inteligence - Total	-0.027	0.043	-0.053	-0.114
Perceiving Emotions	0.011	0.061	0.000	-0.003
Facilitating Thought	0.009	0.023	-0.058	-0.064
Understanding Emotions	-0.089	0.058	-0.105	-0.269 **
Managing Emotions	0.025	0.029	-0.076	-0.066

$N = 142$; * $p < .05$; ** $p < .01$

Figure Captions

Figure 1. Significant psychometric complexity (ψ_C) effects for A) Reasoning Composite by Delay, B) Performance-Avoid Goal Orientations by Outflow, and C) Performance-Prove Goal Orientation by outflow. Note: Here, ψ_C represents the interaction effect between the respective complexity manipulation and the moderator, on performance.

Figure 2. Psychometric learning effects (ψ_L) for A) Verbal reasoning, B) Numerical reasoning, and C) Abstract Reasoning. Note: Here, ψ_L represents the interaction effect between experience (attempt number) and domain-specific reasoning, on performance. Only the Verbal ψ_L was statistically significant.

Figure A1. Decision and feedback interface for Inventory and Workforce Management microworld simulation

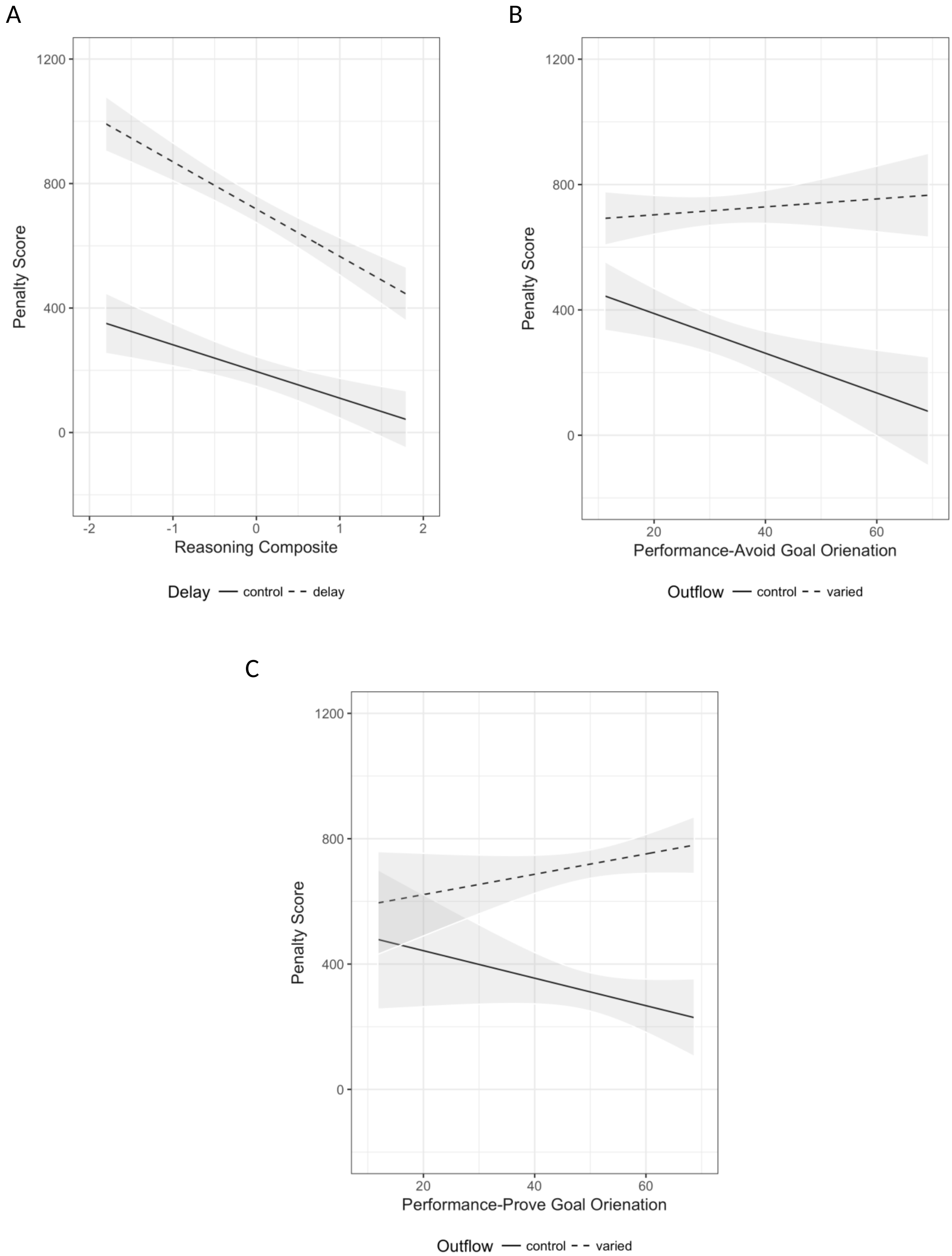

Figure 1. Significant psychometric complexity (ψ_C) effects for A) Reasoning Composite by Delay, B) Performance-Avoid Goal Orientations by Outflow, and C) Performance-Prove Goal Orientation by outflow. Note: Here, ψ_C represents the interaction effect between the respective complexity manipulation and the moderator, on performance.

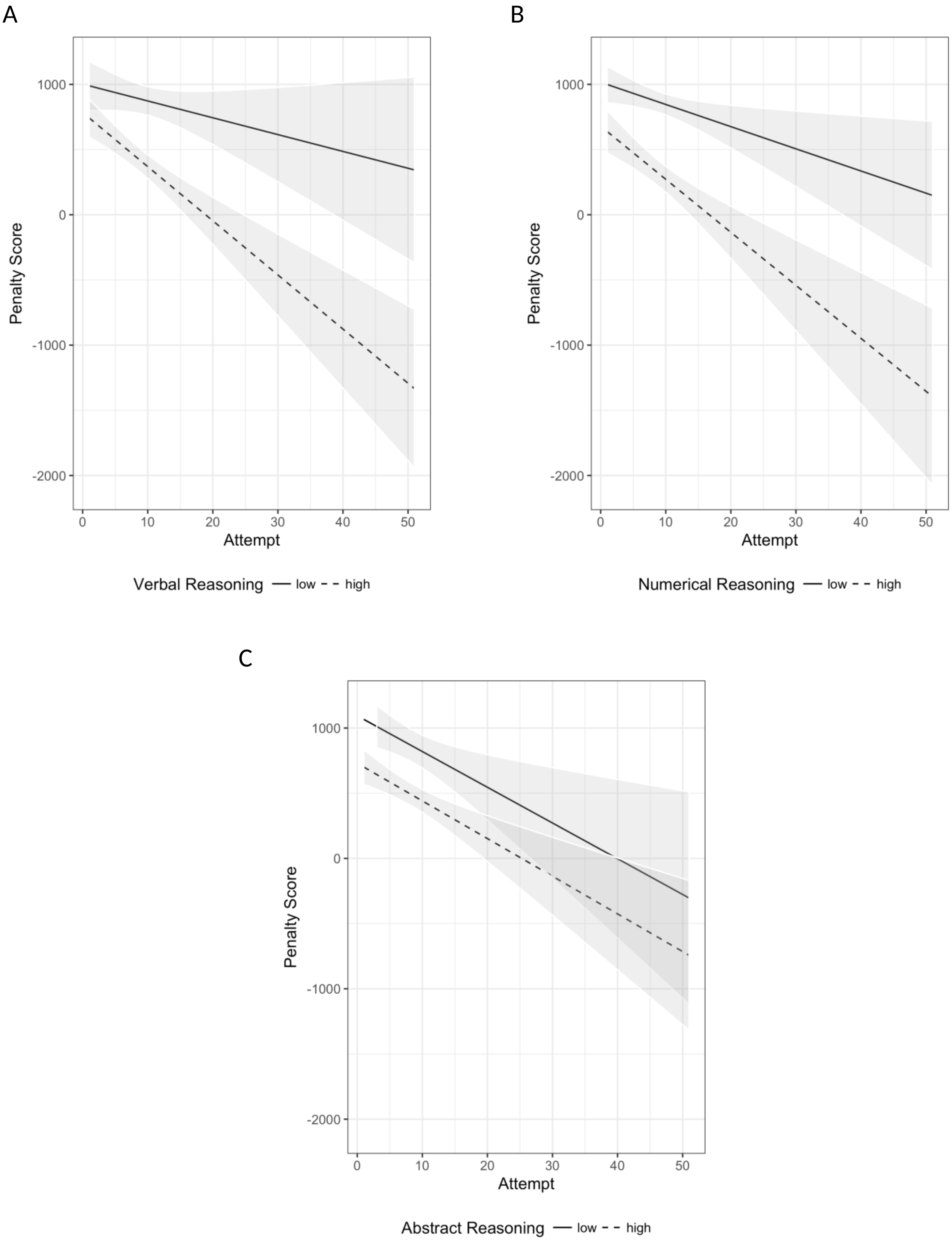

Figure 2. Psychometric learning effects (ψ_L) for A) Verbal reasoning, B) Numerical reasoning, and C) Abstract Reasoning. Note: Here, ψ_L represents the interaction effect between experience (attempt number) and domain-specific reasoning, on performance. Only the Verbal ψ_L was statistically significant.

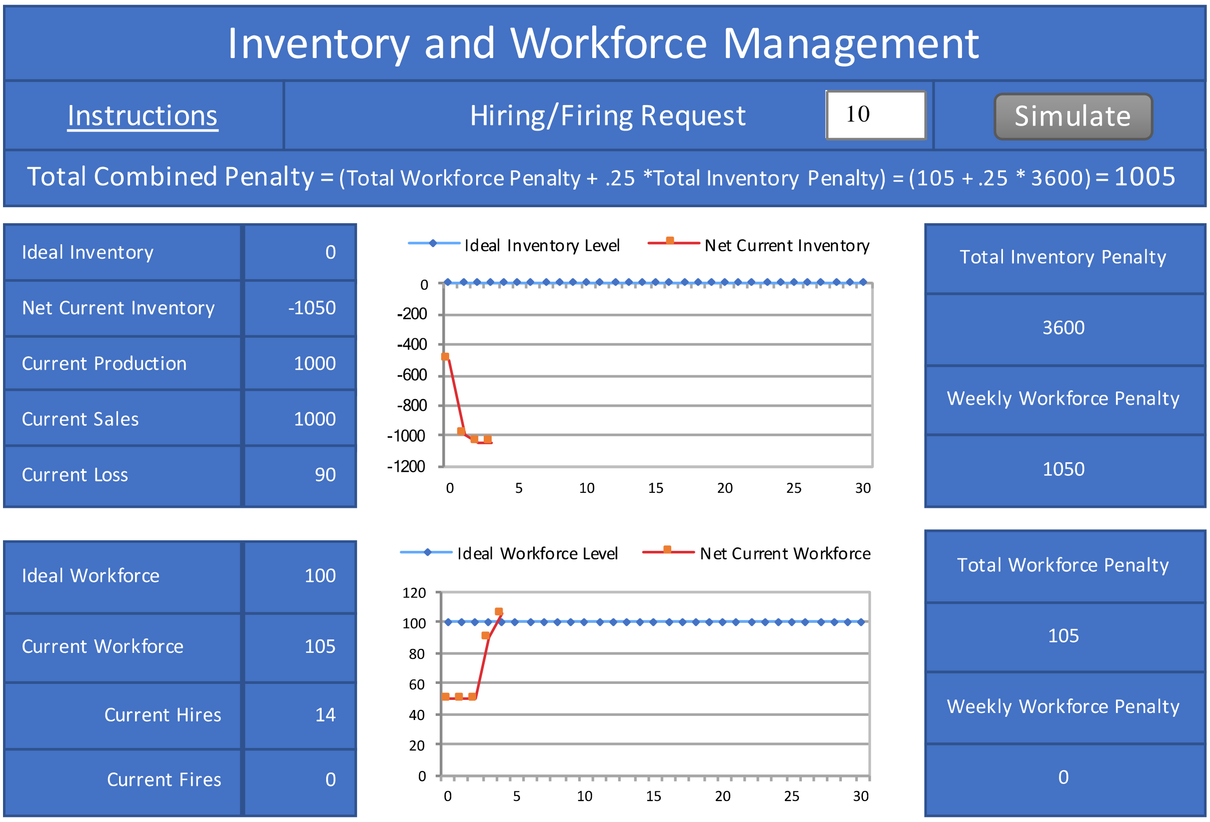

Figure A1. Decision and feedback interface for Inventory and Workforce Management microworld simulation

Appendix A

Inventory and Workforce Management Information Sheet

Background: The goal of this simulation is to introduce you to the impact of lags/delays and outflow associated with inventory/stock management.

You are hired as our new production manager and, now, you are in charge of the production department of our company. Your aim is to manage the finished goods inventory and workforce at their ideal levels.

Our customer demand is quite stable and equal to 1000 items per week. (Note that the amount of sales is equal to the customer demand itself). In order to match the sales, we need to produce 1000 items per week. Each worker produces 10 items per week, so 100 workers are needed to produce the required amount. Therefore, the ideal level of workforce is determined as 100 people. It is for the benefit of the company to maintain the workforce at this constant level, so your managerial skills will be judged by the accumulated discrepancy between the ideal and actual workforce levels.

Delays: You manage the number of workers in your department via hiring/firing requests. Your company can only fire workers by giving notice. In the same way, it takes time for your company to find and hire workers. Therefore, both hiring (positive values) and firing (negative values) requests (your decisions) will be effective after some delay.

Outflow: There may be a loss of stock from the inventory (e.g., through spoiling or by items passing its useable life). This stock outflow might be constant or varied.

Note that the previous manager requested hiring 5 workers per week. Therefore, till your first decision becomes effective, 5 workers will be hired each week according to the initial plan.

Inventory Costs: If we have a positive net inventory, the company pays for storage costs. If we have a negative net inventory, this means we have a backlog (unsatisfied demand), a delay in shipments, which makes our customers unhappy. In the long run, persistent delays in delivery may even result in losing our customers to one of our competitors. Therefore, you need to manage the net inventory too.

The net inventory should be managed at its ideal level, which is equal to zero. Inventory depends both on production and sales. Sales is a constant and it is not a control variable. Therefore, in order to manage the inventory level, you need to control the production, which depends on the workforce. The company does not want to change the productivity level by overtime or underutilization of workers, simply because of the associated high costs. Therefore, you can only control the production via managing the workforce level itself. The ideal situation is to produce the exact amount equal to our sales (i.e. customer demand) keeping the net inventory at zero. We believe that this can be achieved easily because our customer demand is a known constant.

Personal Costs: You will incur a penalty whenever your inventory is above or below zero. This will be another basis for the evaluation of your managerial skills.

The performance evaluation of your company is pretty objective; Your success will be measured based on the following combined penalty formula:

Total Penalty = Total Workforce Penalty + 0.25×Total Inventory Penalty

The previous manager was not a successful one and he generated a total penalty of 3793 for a 30 week period. Therefore, we fired him. Moreover, he left the workforce at a non-ideal level. We hope, you, as our new production manager, will be able to bring the workforce back to its ideal level keeping the inventory at its ideal level too.

We thank you for accepting the position and wish you good luck in your new managerial role.

Simulation Conditions

SM01: There are no lags or any other complex components involved in this version of the game. In this case (outflow = 0), and thus once actual stock = ideal stock, no correction to the stock is required.

SM02: There is a lag effect in this version of the game. The decisions are effective after one week of delay. All else is the same as SM01.

SM03: There is a lag effect in this version of the game. The decisions are effective after three weeks of delay. All else is the same as SM02.

SM04: There is a lag effect in this version of the game. The decisions are effective after five weeks of delay. All else is the same as SM03.

SM05: There are no lags but there is a constant outflow from the stock, which decreases the stock by 10 units per week if no positive correction is made. Thus, in this case (outflow = 10), once actual stock = ideal stock, a correction to the stock is required (so decision must be 10).

SM06: There is a constant outflow from the stock, which decreases the stock by 10 units per week if no positive correction is made. There is a lag effect as well. The decisions are effective after three weeks of delay. All else is the same as SM05.

SM07: There are no lags but there is a random (variable) outflow from the stock, which on the average decreases the stock by 12 units per week if no positive correction is made. Thus, in this case (outflow on-average = 12), once actual stock approaches ideal stock, a correction to the stock is required (so correction must be set up and down around 12 as needed). The main issue here is that an exact value cannot be determined in advance and thus participants must monitor changes to actual stock closely.

SM08: There is a random (variable) outflow from the stock, which on the average decreases the stock by 12 units per week if no positive correction is made. There is a lag effect as well. The decisions (corrections) are effective after three weeks of delay. All else same as SM07.

[Insert Figure A1 about here]

Appendix B

Table B1.
Mixed-level model output (A) cognitive variables, (B) personality variables, (C) Mindset variables, and (D) emotional intelligence.

A	**Baseline**			**Reasoning-Reduced**			**Reasoning**			**Abstract**			**Verbal**			**Numerical**		
	b	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***
Fixed Effects																		
Trials	-0.25	-0.29 | -0.22	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.26 | -0.19	**<.001**	-0.24	-0.27 – -0.20	**<.001**	-0.23	-0.27 – -0.19	**<.001**	-0.23	-0.27 – -0.19	**<.001**
Attempt	-0.24	-0.28 | -0.19	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.37 | -0.24	**<.001**	-0.31	-0.38 – -0.24	**<.001**	-0.31	-0.38 – -0.25	**<.001**	-0.30	-0.37 – -0.23	**<.001**
Outflow	0.17	0.09 | 0.24	**<.001**	0.18	0.10 | 0.27	**<.001**	0.19	0.11 | 0.28	**<.001**	0.20	0.12 – 0.28	**<.001**	0.20	0.12 – 0.28	**<.001**	0.19	0.11 – 0.27	**<.001**
Delay	0.26	0.19 | 0.33	**<.001**	0.25	0.19 | 0.32	**<.001**	0.27	0.20 | 0.33	**<.001**	0.27	0.20 – 0.33	**<.001**	0.27	0.20 – 0.34	**<.001**	0.27	0.20 – 0.33	**<.001**
Outflow x Delay	0.18	0.09 | 0.26	**<.001**	0.17	0.08 | 0.25	**<.001**	0.15	0.06 | 0.24	**<.001**	0.15	0.06 – 0.23	**<.001**	0.15	0.06 – 0.23	**<.001**	0.15	0.06 – 0.24	**<.001**
Reasoning				-0.19	-0.24 | -0.13	**<.001**	-0.17	-0.29 | -0.06	**.004**									
Moderator										-0.09	-0.20 – 0.03	.131	-0.14	-0.26 – -0.03	**.017**	-0.18	-0.29 – -0.06	**.003**
Outflow x Moderator							0.04	-0.05 | 0.14	.382	0.03	-0.07 – 0.12	.580	0.04	-0.06 – 0.14	.414	0.03	-0.06 – 0.13	.492
Delay x Moderator							-0.09	-0.16 | -0.02	**.011**	-0.07	-0.14 – -0.00	**.048**	-0.08	-0.15 – -0.01	**.032**	-0.08	-0.15 – -0.00	**.042**
Attempt x Moderator							-0.05	-0.12 | 0.01	.099	0.00	-0.07 – 0.06	.919	-0.07	-0.13 – -0.00	**.047**	-0.06	-0.12 – 0.01	.076

B	**Neuroticism**			**Openness**			**Agreeableness**			**Conscientiousness**			**Extraversion**		
	b	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***
Fixed Effects															
Trials	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**
Attempt	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**
Outflow	0.19	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**	0.18	0.10 | 0.26	**<.001**
Delay	0.26	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**
Outflow x Delay	0.16	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**
Reasoning	-0.19	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**	-0.18	-0.24 | -0.13	**<.001**	-0.18	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**
Moderator	-0.02	-0.13 | 0.09	.718	0.07	-0.05 | 0.18	.276	0.02	-0.09 | 0.13	.703	0.07	-0.04 | 0.18	.219	0.02	-0.10 | 0.13	.788
Outflow x Moderator	0.02	-0.07 | 0.11	.670	-0.03	-0.13 | 0.06	.498	0.00	-0.09 | 0.10	.936	-0.06	-0.15 | 0.03	.203	-0.01	-0.10 | 0.09	.901
Delay x Moderator	0.00	-0.07 | 0.07	.933	-0.03	-0.10 | 0.05	.482	0.01	-0.06 | 0.08	.744	0.00	-0.07 | 0.07	.922	-0.03	-0.10 | 0.04	.426
Attempt x Moderator	0.01	-0.05 | 0.08	.687	0.03	-0.03 | 0.10	.312	0.00	-0.07 | 0.07	.981	0.04	-0.03 | 0.10	.316	0.02	-0.05 | 0.09	.519

Table B1 (continued).
Mixed-level model output (A) cognitive variables, (B) personality variables, (C) Mindset variables, and (D) emotional intelligence.

C	**Learning Goal**			**Performance-Avoid**			**Performance-Prove**			**Need for Cognition**			**Implicit Theories - Intell.**			**Implicit Theories - Pers.**		
	b	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***
Fixed Effects																		
Trials	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**
Attempt	-0.3	-0.37 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.3	-0.37 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**
Outflow	0.18	0.10 | 0.26	**<.001**	0.19	0.10 | 0.27	**<.001**	0.18	0.10 | 0.26	**<.001**	0.19	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**
Delay	0.25	0.19 | 0.32	**<.001**	0.26	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.26	0.19 | 0.32	**<.001**	0.26	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**
Outflow x Delay	0.17	0.08 | 0.25	**<.001**	0.16	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**	0.16	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**
Reasoning	-0.19	-0.24 | -0.13	**<.001**	-0.18	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**	-0.20	-0.26 | -0.14	**<.001**	-0.18	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**
Moderator	0.07	-0.03 | 0.18	.176	-0.14	-0.25 | -0.03	**.014**	-0.11	-0.22 | 0.00	.063	0.11	-0.00 | 0.23	.060	0.02	-0.10 | 0.13	.783	0.00	-0.11 | 0.11	.950
Outflow x Moderator	-0.06	-0.15 | 0.03	.165	0.13	0.04 | 0.22	**.005**	0.11	0.02 | 0.20	**.024**	-0.03	-0.12 | 0.06	.533	0.01	-0.08 | 0.10	.818	0.00	-0.09 | 0.09	.998
Delay x Moderator	-0.01	-0.08 | 0.06	.848	0.00	-0.07 | 0.07	.908	0.04	-0.03 | 0.11	.294	-0.07	-0.14 | 0.00	.059	0.04	-0.03 | 0.11	.245	0.00	-0.07 | 0.07	.983
Attempt x Moderator	-0.02	-0.08 | 0.05	.572	-0.02	-0.09 | 0.04	.495	-0.05	-0.12 | 0.01	.096	-0.01	-0.08 | 0.06	.790	0.02	-0.05 | 0.08	.644	0.01	-0.06 | 0.07	.802

D	**Emotional Intelligence (total)**			**Understanding Emotions**			**Managing Emotions**			**Perceiving Emotions**			**Facilitating Thought**		
	b	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***	***b***	***CI***	***p***
Fixed Effects															
Trials	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**	-0.23	-0.27 | -0.19	**<.001**
Attempt	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.37 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**	-0.31	-0.38 | -0.24	**<.001**
Outflow	0.18	0.10 | 0.27	**<.001**	0.19	0.10 | 0.27	**<.001**	0.19	0.10 | 0.27	**<.001**	0.19	0.10 | 0.27	**<.001**	0.18	0.10 | 0.27	**<.001**
Delay	0.25	0.19 | 0.32	**<.001**	0.26	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**	0.25	0.19 | 0.32	**<.001**
Outflow x Delay	0.16	0.08 | 0.25	**<.001**	0.16	0.08 | 0.25	**<.001**	0.16	0.08 | 0.25	**<.001**	0.16	0.08 | 0.25	**<.001**	0.17	0.08 | 0.25	**<.001**
Reasoning	-0.18	-0.24 | -0.13	**<.001**	-0.18	-0.24 | -0.12	**<.001**	-0.18	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**	-0.19	-0.24 | -0.13	**<.001**
Moderator	-0.01	-0.13 | 0.11	.861	0.05	-0.06 | 0.17	.380	-0.01	-0.12 | 0.11	.926	-0.03	-0.14 | 0.08	.579	0.01	-0.11 | 0.12	.883
Outflow x Moderator	-0.03	-0.12 | 0.07	.608	-0.04	-0.14 | 0.05	.377	-0.04	-0.14 | 0.05	.345	0	-0.10 | 0.09	.972	-0.03	-0.12 | 0.07	.548
Delay x Moderator	0.02	-0.05 | 0.09	.581	-0.06	-0.13 | 0.01	.095	0.04	-0.03 | 0.11	.314	0.04	-0.03 | 0.11	.257	0.02	-0.05 | 0.09	.626
Attempt x Moderator	0.03	-0.04 | 0.10	.446	0.03	-0.03 | 0.10	.318	0.02	-0.05 | 0.09	.578	0.03	-0.04 | 0.11	.357	0.01	-0.05 | 0.08	.755

Appendix C

Statement of Data Transparency

The research group conducted expertise research and provided a 2-year leadership training program for mid-level managers from large organizations from 2006-2015. The assessment and professional development component is a core feature of the program. It has a theory based, elaborated assessment-for-learning focus and is structured around five related components: 1) objective (psychometric) assessment on cognitive and non-cognitive constructs relevant to work settings; (2) instruction in the details the very same constructs; 3) individualized feedback and guidelines for self-development activities contingent on this information; 4) professional instruction in the significance of the constructs in different work roles; and 5) exposure to practical experience in the lab and back-on-the-job. The objective of this approach was to foster the development of flexible expertise in managerial leadership that extends beyond domain-specific routine expertise. Data for a variety of experiments and learning activities were collected from participating managers across the two-year period. The total ALL participant pool consists of a core cohort of 423 industry managers and individuals associated in some way with the managers (including, the supervisors of n = 152 managers, and n = 1266 peers/associates of the managers). A second cohort consists of ~ 5000 student and community participants unrelated to the manager cohort.

The data reported in this manuscript have been previously published and, as described above, were collected as part of a larger data collection across multiple points in time. Findings from the manager cohort data collection have been reported in separate manuscripts. MS01 (the current manuscript) focuses on variable 12.1 and variables from set 4, 5, 6 and 10. MS02 (published) focuses on variable 12.3. MS03 (published) focuses on variable set 7 and 4.9. MS04 (published) focuses on 6.1 versus 7.1 and 6.5 versus 7.5, and controls for variable 4.1. The core focus of MS05 (published) is an administration of variable 4.1 under different conditions, and its relationship with 4.2 – 4.4 and the set 6 variables. MS06 (published) focus was on variable 2.2 and also considered variables 1.4 and 1.5 specifically. MS07 (published) was a validation study of variable 4.5 and 4.6 in relation to variables 4.1 – 4.4. MS08 (published) considered the relationship between 6.5 and 7.5 across different occasions and evaluated the impact of occasion-specific perceptions (additional measures associated with 6.5 and 6.7), it also considered variables 5.3 and 4.10. MS09 (published) extended on the methodology of MS08 by considering variables 8.1 and 8.2 in field settings (rather than in the lab) in relation to specific sub-facets of 5.6 and 5.7. MS10 (under review) focused on 2.1 and its relationship with variables 6.1, 6.5, 7.1 and 7.2. MS11 (under review) focused on 6.1, 6.2 and 6.5 and sub-facets of 8.1, 8.2 and 8.5 assessed across different occasions, while also considering variables 10.2 – 10.4. MS12 (published) focused in depth on 7.5 under a range of different situations. MS13 (under review) focuses on variable set 11 and relationships with 5.5 – 5.7. MS14 (planned) will focus on variable set 9.

Table C1.

Manuscripts published, under-review and planned.

Table C1		**MS 01**	**MS 02**	**MS 03**	**MS 04**	**MS 05**	**MS 06**	**MS 07**	**MS 08**	**MS 09**	**MS 10**	**MS 11**	**MS 12**	**MS 13**	**MS 14**
Participant Pool N=423		142	71	121	115	252	172	264	132	135	131	201	111	256	TBD
Demographics															
1.1	Age														
1.2	Sex														
1.3	Education														
1.4	Years in current job						x								
1.5	Years management experience						x								
Job Performance															
2.1	Supervisor job-performance ratings										x				
2.2	Concept Evaluation Survey						x								
2.3	Strategy Talk Evaluations														
Personal Projects															
3.1	Identity Relevance														
3.2	Value Congruence														
Cognitive Ability															
4.1	RPM				x	x		x							
4.2	Verbal Reasoning	x				x		x							
4.3	Numeric Reasoning	x				x		x							
4.4	Abstract Reasoning	x				x		x							
4.5	Latin Square Task							x							
4.6	Greco Latin Square Task							x							
4.7	Counter-Factual Analogies														

Table C1		MS 01	MS 02	MS 03	MS 04	MS 05	MS 06	MS 07	MS 08	MS 09	MS 10	MS 11	MS 12	MS 13	MS 14
4.8	Premised Figural Analogies														
4.9	Modified Synthesis Task			x											
4.10	Series completion								x			x			
4.11	Flexible Inference														
4.12	Flexible Mapping														
Motivation															
5.1	Implicit Theories - Intelligence	x													
5.2	Implicit Theories -Personality	x													
5.3	Need for Cognition	x							x						
5.4	Need for Novelty														
5.5	Goal Orientation: Learning													x	
5.6	Goal Orientation: Performance Prove									x				x	
5.7	Goal Orientation: Performance Avoid									x				x	
Personality (Trait)															
6.1	Neuroticism	x			x	x					x	x			
6.2	Extraversion	x				x						x			
6.3	Openness	x				x									
6.4	Agreeableness	x				x									
6.5	Conscientiousness	x			x	x			x		x	x			
Personality (State, lab-based-ESM; + occasion-specific perceptions)															
7.1	Neuroticism			x	x						x1				
7.2	Extraversion			x							x1				
7.3	Openness			x											
7.4	Agreeableness			x											

Table C1		MS 01	MS 02	MS 03	MS 04	MS 05	MS 06	MS 07	MS 08	MS 09	MS 10	MS 11	MS 12	MS 13	MS 14
7.5	Conscientiousness			x	x				x				x1		
Personality (State, field-based-ESM; + occasion-specific perceptions)															
8.1	Neuroticism									x1		x1			
8.2	Extraversion									x1		x1			
8.3	Openness														
8.4	Agreeableness														
8.5	Conscientiousness											x1			
Personality (Frames of Reference)															
9.1	Neuroticism														x
9.2	Extraversion														x
9.3	Openness														x
9.4	Agreeableness														x
9.5	Conscientiousness														x
Emotional Intelligence															
10.1	PE1	x													
10.2	FT1	x										x			
10.3	UE1	x										x			
10.4	ME1	x										x			
Thinking Styles															
11.1	Legislative													x	
11.2	Executive													x	
11.3	Judicial													x	
11.4	Hierarchical													x	
11.5	Monarchic													x	
11.6	Oligarchic													x	

Table C1		MS 01	MS 02	MS 03	MS 04	MS 05	MS 06	MS 07	MS 08	MS 09	MS 10	MS 11	MS 12	MS 13	MS 14
11.7	Anarchic													x	
11.8	Global													x	
11.9	Local													x	
11.10	Internal													x	
11.11	External													x	
11.12	Liberal													x	
11.13	Conservative													x	
Simulations/Microworlds															
12.1	Inventory Management	x													
12.2	Employee Management (Team)														
12.3	Employee Management (Individual)		x												
12.4	New Product Launch														
12.5	Complex Problem Solving														

Notes: x1 = sub-facets of the variable are considered; MS01 = current manuscript; MS10, MS11, and MS13 are under review; MS14 is a planned future publication; MS02 – MS09 and MS12 are published. Demographic variables (set 1) are reported in all papers, only those publications indicated have specifically considered them as design variables.