Exploring Ways to Measure Entrepreneurial Mindset: The development of a Student-Focused Effectual Logic Assessment Instrument

Mr. Todd Mathew Fernandez, Purdue University, West Lafayette (College of Engineering)

Todd is a PhD Candidate in Engineering Education at Purdue University who’s research is focused on entrepreneurship education and entrepreneurship education as a component of modern engineering education efforts.

Dr. Nathalie Duval-Couetil, Purdue University, West Lafayette (College of Engineering)

Nathalie Duval-Couetil is the Director of the Certificate in Entrepreneurship and Innovation Program, Associate Director of the Burton D. Morgan Center, and an Associate Professor in the Department of Technology Leadership and Innovation at Purdue University. She is responsible for the launch and development of the university’s multidisciplinary undergraduate entrepreneurship program, which has involved over 6000 students from all majors since 2005. She has established entrepreneurship capstone, global entrepreneurship, and women and leadership courses and initiatives at both the undergraduate and graduate levels. Prior to her work in academia, Nathalie spent several years in the field of market research and business strategy consulting in Europe and the United States with Booz Allen and Hamilton and Data and Strategies Group. She received a BA from the University of Massachusetts at Amherst, an MBA from Babson College, and MS and PhD degrees from Purdue University. She currently serves on the board of the United States Association for Small Business and Entrepreneurship in the role of Vice President for Research. She is also a Senior Research Advisor to the Stanford University Epicenter.

©American Society for Engineering Education, 2017
Exploring Ways to Measure Entrepreneurial Mindset: The Development of a Student-focused Effectual Logic Assessment Instrument

Introduction
The theory of effectuation proposes a model of entrepreneurial decision making that explains how expert entrepreneurs pursue new venture creation in non-predictive ways. The basis of the theory, an ‘effectual logic’ is demonstrated through a set of heuristics that drive the behavior of experienced and successful entrepreneurs. In her initial proposal of effectuation, Sarasvathy identified the theory as valuable content to be integrated into entrepreneurship education given that it is an explanation of entrepreneurial action grounded in the behavior of real entrepreneurs.

Since the proposal to include effectuation in entrepreneurship education, there has been a growth in efforts to include effectuation through teaching and educational research. These attempts include to measure or assess students’ use of effectuation have included classroom observations, discussion questions, verbal protocols, and several multiple choice style instruments. The efforts to develop effectuation as education have also included textbooks and instructional materials available from the ‘Society for Effectual Action’. While each of these pieces of scholarship approaches effectuation education differently, they all provide an increasing body of knowledge on which to teach effectuation. Most relevant to our work has been the ongoing development of the survey instruments which seek ways to quantify and measure effectual behavior for research and teaching.

However, the instruments and methods designed to measure effectuation are collectively limited in two critical ways. First, they perpetuate a problem noted in the entrepreneurship education literature of not employing established educational approaches to validate their use with students. Second, and more problematic for measuring effectuation, they do not explore the situated nature of entrepreneurial cognition as described in effectuation theory and research. Rather than assessing how behavior varies among differing contexts or tying changes in thinking to specific experiences, existing effectuation measurement techniques simply look to see if the student applies the ‘correct’ knowledge and extrapolates that they do so for appropriate reasons. As such, these instruments do not explore the metacognitive models that Haynie et al. explain are necessary. These models are critical to the underlying conception of uncertainty, and applicability of different cognitive approaches, that are at the heart of effectuation.

In this study, we explore the development of an instrument, referred to as the Factorial Effectuation Survey (FES) that addresses both of those gaps. FES will measure whether and how students’ decision making approaches change as the context in which those decisions occur changes. More specifically, the instrument will use a group of questions that are repeated across multiple scenarios (i.e., contexts). Students will select from choices representing operationalization of causal or effectual heuristics for a given scenario. The scenarios are designed to differentiated student behavior on dimensions that are theorized as appropriate (e.g., the type of uncertainty) and inappropriate (e.g., the size of the business) reasons to pursue effectual or causal approaches.

FES is targeted for use in engineering education and with engineering students. The instrument uses a novel construction that employs the factorial survey technique to empirically measure
how students situate their decision making, rather than simply exploring their cognition without attention to context or within a single, typically entrepreneurial, situation, about which results are generalized. This paper details the instrument’s intended use, a pilot study of the instruments’ structure, sensitivity, and reliability; and further work that needs to be undertaken to create an instrument able to guide and inform the teaching of effectuation in the classroom.

**Theoretical framework**

The theory of effectuation was first proposed by Sarasvathy in the early 2000’s as an explanation for how expert entrepreneurs act when faced with decisions at the early stage of venture creation. The work takes a grounded theory approach to unpacking the entrepreneurial behavior termed ‘intuitive’ by Knight in 1921. The intuition that Knight names, and which Sarasvathy explains, describes the behavior of individuals when faced with highly uncertain situations where it is impossible to make meaningful calculations of risk. Sarasvathy proposed that expert entrepreneurs eschew methods that predict the likelihood of certain events (i.e., predictive approaches), and instead focused on pursuing opportunities where they are able to exert a higher level of control on the outcome (i.e., effectual approaches).

Effectuation builds on work by Simon and others, proposing the bounded rational model of human cognition and identifying heuristics that humans use to make decisions. Sarasvathy proposed five heuristics that represent thinking in the effectual style of logic and situates them in opposition to what are referred to as ‘causal’ heuristics that aid decisions through predictive approaches and planning. The two sets of heuristics are reproduced in Table 1 below.

<table>
<thead>
<tr>
<th>Situation / Issue</th>
<th>Causal Approach</th>
<th>Effectual Heuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>View of the future</td>
<td>Define goals and gather resources to support those goals</td>
<td>Assess what you have on hand, use those resources to pick a viable goal</td>
</tr>
<tr>
<td>Tolerance of risk</td>
<td>Focus on what has the highest return for a given calculation of risk</td>
<td>Focus on taking actions that only risk what you can afford to lose</td>
</tr>
<tr>
<td>Approach to others</td>
<td>Analyze competitors for their weaknesses</td>
<td>Partner with potential competitors to learn and understand</td>
</tr>
<tr>
<td>View of the unexpected</td>
<td>To be avoided, shows a lack of proper planning or preparation</td>
<td>To be learned from, highlights opportunities to learn and adjust</td>
</tr>
<tr>
<td>Basis for taking action</td>
<td>Focus on what you can predict</td>
<td>Focus on what you can influence</td>
</tr>
</tbody>
</table>

The heuristics represent an inner core of the relationship between effectuation research and a larger context of decision making. The layers represent a series of hierarchical levels in which elements of Sarasvathy’s explanation exist and their order with each inner layer containing a non-exclusive or complete set of techniques for addressing the challenges that arise from the next outer layer. At the center, effectual heuristics represent an inherently incomplete way of implementing a broader framework of decision making. Therefore, the effectual heuristics...
represent an incomplete set of implementations of effectuation theory. The theory of effectuation contains the complete, but theoretical, set of heuristics used to solve entrepreneurial problems in using a non-predictive logical approach. The theory may contain more heuristics, as expected in work describing heuristic approaches to explaining cognition\textsuperscript{16,17}, but exists definitely within a class of solution techniques based on non-predictive logic\textsuperscript{19}. That non-predictive logic includes the theory of effectuation, but may or may not contain other approaches as well. Reaching the boundary to the outer layer, non-predictive approaches are one theorized explanation of decision making approaches to solve problems with Knightian uncertainty. However, while this is one proposed explanation, nothing in Sarasvathy’s work proposes the effectual heuristics, Theory of effectuation, or non-predictive logic approaches as the sole approach to addressing Knightian Uncertainty.

\textbf{Figure 1 Layering of effectuation theory and effectuation concepts within business strategy}

To be specific, the FES instrument measures the use of effectual heuristics with a goal of generalizing those measurements to a broader discussion of the Theory of Effectuation. While further generalization might occur, such as through correlations with other instruments, the contextualization in business decisions limits generalization of results to the Theory of Effectuation level because that is the outer most level which is specific to business decisions.

\textbf{Intended use and hypotheses}
The development process that we have used to create FES is informed by three existing bodies of work:

(1) Conceptualizations of validity in assessment development that focus on use and interpretation as the core of validity\textsuperscript{20,21}
Established processes for developing educational assessment instruments and factorial surveys.

Existing research on effectuation theory comprising data collection on students and experts.

The bodies of work each focus on a different element of the instrument development process. Collectively, they provide a framework for what a good instrument includes (the goal), what process are used to achieve that framework (the process), and the theory that is operationalized by the development process (the input). They are not ordered or hierarchical, but rather three legs in a metaphorical stool that should be seen as working together to support FES’ use.

**Defining how we approach validity**

The conceptualization of validity used in this study builds primarily on work by Messick and Kane. It consists of two primary components: 1) what validity is, and 2) how validity is assessed. While the concepts are broadly applicable to many types of measurement and assessment, the language used herein is specific to the application of these concepts to FES.

Messick’s work focuses on refining the conceptualization of what validity is. He argues that the conception of measurement instruments as having a set of multiple validities that are treated independently is outdated. Instead, he proposes a singular, unified validity, but that validity contains multiple aspects. This differs from prior conceptions of validity that discuss multiple validities separately (e.g., content validity or criterion validity). The problem with the multiple validities approach is its relationship to the core concept of studying validity, which is evaluative of more than just the instrument’s properties. Instead of focusing on the properties of the instrument, these conceptions focus on how the scores will be used, and what evidence supports that use. A student might complete a thoroughly developed test, but if their score is used to make decisions that are not related to the test’s content, then that use would likely be considered invalid.

Kane adds a clear framework for how Messick’s unified validity is assessed or supported. He approaches supporting validity as an argument, rather than an assessment. Kane draws on Popper’s stages of theory development where theories are proposed and then subject to attempts to refute them. Kane proposes that arguments for validity generally involve two stages: 1) The formative, in which an argument for validity is planned and articulated, and 2) the summative stage, in which the argument is tested. This terminology intentionally parallels terminology from program evaluation work. In effect, the approach to validity is an operationalization of theory and practice to evaluate whether an instrument does what it claims to do. He formulates the proposal of an argument for validity around the concept of a logical chain with multiple different types and pieces of evidence working together as different ‘links’. The chain approach begins with a series of conjectures. These conjectures form an interpretive argument for why the instrument may be interpreted in a certain way to serve a proposed purpose. The sequence of conjectures forms the interpretive argument. The conjectures are then subject to a series of tests to assess their viability or refutability. The results of those studies of the interpretive arguments’ conjectures forms the validity argument.

In keeping with Potter’s approach, the studies which form the validity argument typically propose and seek to refute a proposed hypothesis rather than propose and answer research
questions. This is because theory form a Potterian perspective, is not found but proposed and then accepted until it is refuted. Evidence is collected that assess the proposed hypothesis as well as refute alternative hypotheses that the test developers have identified. This includes not just the assertions that the developers make, but the assumptions that are either explicit or implicit in those assumptions as well.

The argument for validity is inclusive of all aspects of the development of a test for a specific use. The argumentation chain often includes evidence that formal processes for developing the instrument were used and that the processes are grounded in best practices. Other support for the validity argument can include theoretical justifications, best practices or other guidance from existing literature, or even appeals and assertions of common sense. The assessment of the validity argument and the assessment of the interpretive argument, made via validity evidence, occurs holistically from an evaluation of whether the proposed interpretation is the most likely or reasonably hypothesis. To operationalize this approach, the sections below detail the intended use of FES, propose our interpretive and validity arguments, and detail work to date. We end this section by proposing the specific hypotheses that we will test in this study.

**Intended use**
FES focuses on measuring predictors of students’ selection of effectual or causal approaches to business problem solving. The intended use of FES rises from gaps in the measurement of effectuation education. We specifically target that use at engineering education and engineering students given that there is a strong movement to equip engineering students with business skills, often framed through entrepreneurship, as it provides a means to connect technical engineering skills to the creation of value. A gap that must be addressed is that movement has occurred without a well-aligned set of assessment instruments to support our understanding of engineering student progression in these areas.

The use of effectuation and causation as a framework for engineering entrepreneurship education is particularly pertinent. Engineering epistemologies call for both technical and design skills in engineers. These generally match the underlying logic of decision making present in effectual and causal logic. Students are heavily trained through their engineering courses in predictive (i.e., causal) ways of making decisions. Therefore, a tool such as FES which can measure how engineering students employ or do not employ that predictive basis in a new environment of entrepreneurship, where it may not be appropriate, is highly useful.

The interpretation of these measures, involves the way sense is made of the raw measurements. In this case, the instrument, by utilizing the intentional variation of context, can be interpreted as illustrating what (or whether) variation in context influences students’ selection of causal and effectual heuristics. The intentional variation of context on specific dimensions that, theoretically, should or should not affect students’ approaches is based on work for measuring social judgements in sociology.

The use of that interpretation is to develop a greater understanding of how students use effectuation and how that use shifts over time or education. The first, and primary, use of the interpretation will aid educators in understanding what predicts the use of effectuation by students and whether there is evidence that students use effectuation in a situated way (i.e., given a particular context or scenario) as seen in experts and as proposed by Knight. The use of the
measurement and interpretation in this way would be operationalized through the design of curricular interventions design to teach the proper (i.e., contextually appropriate) use of effectuation as well as target interventions to certain groups who observed to over or under use effectuation overall. The secondary use is to track whether students are developing the capacity to appropriately situate the use of effectuation and causation. By tracking the progression of individual or groups of students in a pre-post test design, the interpretation will allow instructors to assess whether students’ use of effectual and causal logic is becoming more appropriately situated. While this is not the primary use being validated in this study, it is based on the same evidence. The use as a developmental tracker extends the use from simply identifying gaps to targeting whether those gaps are closed.

*Interpretive argument and this study’s purpose*

To support the use I described above, we propose the following interpretive argument. Kane\(^{25}\) highlights three criterion (i.e., characteristics) of a convincing ‘practical’ argument: 1) clarity of argument, 2) coherence of argument, and 3) the plausibility of the assumptions in the argument, which he extends to interpretive and validity argument. The proposed interpretive argument is shown in Figure 2 on the next page. The chevrons represent dependent levels of the argument (i.e., aspects of validity) with each box containing a supporting assertion. To be clear, this argument does not propose a validity argument but proposes an interpretive argument for which we will develop the supporting validity argument over a series of studies.

Because of the extent of evidence required to create the validity argument from the interpretive argument shown below, we are approaching this effort as a series of studies and efforts which individually look at different assertions within the argument. The study presented in this paper serves as one of those studies. Specifically, this study addresses an initial assessment of the theoretical and criterion components of the interpretive argument as well as providing an overall pilot test of the items within FES.
Figure 2. Proposed interpretive argument and proposed evidence to construct the validity argument
**Hypotheses of this study**

As noted in the previous section, a hypothesis driven approach is being employed to study the validity of FES for the planned use. This study specifically addresses 3 hypotheses:

1. The items in FES function as a coherent group to measure the intended trait.
2. All of the items in FES will provide useful information for separating students’ use of effectual and causal logic.
3. Student behavior on the FES measure will be stable across time at the individual and group level when students are not instructed in effectual, entrepreneurial, or business techniques.

**Instrument**

FES will measure whether students’ decision making approaches change as the context in which those decisions changes. More specifically, the instrument will use a group of questions that are repeated across multiple decision scenarios (i.e., contexts). Students select from choices representing operationalization of causal or effectual heuristics for a given scenario. The scenarios are designed to differentiated student behavior on dimensions that are theorized as appropriate (e.g., the type of uncertainty) and inappropriate (e.g., the size of the business) reasons to pursue effectual or causal approaches. Combined with other predictors, such as demographics and entrepreneurial experience or education, this enables the measurements to be used to interpret whether students.

The instrument builds on the factorial survey approach to studying the influence of context on social judgements. Using the factorial approach, test takers respond to the same item or items multiple times in a variety of contexts. The contexts are intentionally varied on a set of dimensions (e.g., the age of a patient or the gender of a boss) to create multiple predefined scenarios. Because the questions are identical for all scenarios, the variation in a test taker’s response provides evidence, on a given dimension of context affected, of that individual’s decision making. At scale, these judgements provide an understanding of what societal models for cognition exist in a quasi-experimental way.

In this case, the implementation of a factorial survey for FES will be a binary choice selected response item which students complete electronically. Participants will complete 4 scenarios that vary on 2 dimensions. In each scenario, the test takers respond to 5 questions situated within that scenario. For this paper, we demonstrate this through a simplified FES involving only 1 scenario and the 5 questions mentioned above. The goal of that simplification is to first assess the performance of individual items before examining performance of the multiscenario version of FES.

**Instrument development process**

Despite calls since early in the development of the factorial survey approach, limited information exists proposing and supporting processes for rigorous creation of such instruments. However, there do exist general guidelines and practical suggestions for development that can be drawn from literature on prior implementations of the factorial survey. To supplement the practical suggestions, we employed work on test item development to guide the development of the items, which are generally treated as secondary efforts in reporting on factorial survey development. Building on those studies, we plan to use the development process shown in Figure 3. This paper specifically focuses on the steps outlined in red within that process.
The process initially began with the identification of the measured trait that we will use in the instrument. From that trait we will define a master scenario and a set of initial items.

**Measured trait**
For the measured trait, we decided to use definitions and interpretations of effectuation, and the effectual heuristics, that are as close to those created by Sarasvathy as possible. For those definitions, we used the most recent publication of basic effectuation information as posted on the website of the society for effectual action, run by the original researchers. The body of research on effectuation is generally consistent and intentional about making use of the originally proposed definition. Use of the most current work published by the original researchers is appropriate because the original theory developers are actively engaged in continuing to develop the theory, and the theory is in an emergent stage. Therefore, we adopted and will operationalize the definition of effectuation as:

*Effectuation is a logic of thinking that uniquely serves entrepreneurs in starting a business [and] provides a way to control a future that is inherently unpredictable*

We interpret that definition as characterizing effectuation through three primary characteristics: 1) An approach to thinking that focuses on control, 2) That that approach is defined in relative opposition to predictive approaches to problem solving, and 3) that it is appropriately used when situated in contexts characterized by the lack of ability to meaningfully predict outcomes.

Further, we define the effectual heuristics and their causal contrasts from Sarasvathy (2011) as follows:
The first heuristic is:

**Start with your means** - When expert entrepreneurs set out to build a new venture, they start with their means: who I am, what I know, and whom I know. Then, the entrepreneurs imagine possibilities that originate from their means.

Start with your means contrasts with:

**Pre-set goals or opportunities** - Causal reasoning works inversely by assembling means after a goal is set.

The second heuristic is:

**Focus on the downside risk** - Expert entrepreneurs limit risk by understanding what they can afford to lose at each step, instead of seeking large all-or-nothing opportunities. They choose goals and actions where there is upside even if the downside ends up happening.

Focus on the downside risks contrasts with:

**Expected return** - Causal reasoning first targets a return, then works to minimize associated risk.

The third heuristics is:

**Leverage contingencies** - Expert entrepreneurs invite the surprise factor. Instead of making “what-if” scenarios to deal with worst-case scenarios, experts interpret “bad” news and surprises as potential clues to create new markets.

Leverage contingencies contrasts with:

**Avoiding surprises** - Causal reasoning works to minimize the probability of unexpected outcomes.

The fourth heuristic is:

**Form partnerships** - Expert entrepreneurs build partnerships with self-selecting stakeholders. By obtaining pre-commitments from these key partners early on in the venture, experts reduce uncertainty and co-create the new market with its interested participants.

Form partnerships contrasts with:

**Competitive analysis** - Causal reasoning presumes that competitors are rivals to contend with.
The fifth heuristic is:

**Control vs. Predict** - By focusing on activities within their control, expert entrepreneurs know their actions will result in the desired outcomes. An effectual worldview is rooted in the belief that the future is neither found nor predicted, but rather made.

Control vs. predict contrasts with:

**Inevitable trends** - Causal reasoning accepts that established market forces will cause the future unfold.

The definition of effectuation and the heuristics were then used to develop the master scenario and the test items described in Figure 3. For the initial version of the instrument we chose to develop a master scenario and items that mirror theoretical definitions as closely as possible.

**Master scenario**
Following the process that we described above, the master scenario was the first portion of the instrument that we developed. The master scenario defines the broad context in which test takers will act. It establishes the scope of the situation, the particulars, and what details are or are not provided to the participant.

For FES, we chose to use a scenario that closely matched the context used to situate participants responses in Sarasvathy’s original study. Sarasvathy used a verbal protocol that placed participants as founders of an educational software venture, with a product, but with no customers or market yet served.

From that initial context, we refined the language slightly to make it more readable for the student audience. We also began to identify dimensions on which we planned to vary the master scenario and used that process iteratively to help define what information the scenarios will provide, and how that information will vary in different scenarios. The final master scenario that we used for this study is:

You have created an entrepreneurship simulator video game called ‘Venturing Inc’. You believe that it could be used both for entertainment and entrepreneurship education. You have acted to protect your intellectual property, and you have filed paperwork to incorporate the company under the name Entrepreneurship, Inc. However, you have not yet begun to move forward on identifying customers or bringing the product to market. In sum, you have created an initial product, and have taken legal steps to begin the business, but have not moved forward in starting the business yet.
Item development

Following the creation of the master scenario, we used the definitions of the heuristics noted above to write a series of items. While other survey approaches to studying effectuation\textsuperscript{4,6,37} have written many items and then reduced that to a reasonable number via studies testing item clarity, function, and structure, the use of the factorial survey approach necessitates a smaller number of more focused initial items. Most factorial surveys usually take this approach to the extreme and use only one item, which is repeated the only observed variable for each scenario. Given the small body of research on the individual heuristics, how they interrelate in practice, and how they manifest in students (Fernandez & Duval-Couetil, 2017) we decided that an appropriate middle ground was to write a single item for each heuristic, totaling five items. This serves both the appropriate levels of content coverage\textsuperscript{22} as well as the efficiency of item use necessary when test takers complete the items multiple times across the scenarios.

For the initial version of each item, we choose to take the most direct approach to operationalizing the definitions possible. We first wrote a question stem representative of each ‘issue’ from Table 1. The stems took the form of a partial sentence that is completed by either an effectual or causal response. For the binary responses, we reworking the effectual and causal heuristic definitions from Table 1 to form the second half of the sentence. An example of one item, the item for the ‘approach to others’ decisions, appears below:

When identifying early customers and partners I would...

A) focus on analyzing my competitors so that I can identify their weaknesses.

B) engage with people I trust and build partnerships even if they might later become rivals.

The major changes in wording between the heuristic definition and items came from reframing the descriptions provided of the heuristics into actions that might be taken by an individual test taker. Throughout, for the initial item development and testing, we sought to be as authentic as possible to the original wording of the trait definitions provided stated by Sarasvathy (2008). The final versions of all the initial items, which we used in this study, appear in the appendix.

Methodology

Population and data collection

We collected data for this study from a required first semester course at the [blinded] University College of Business. The course generally covers an introduction to business, and specifically focuses on helping students understand the difference sub-disciplines within management. The course does not cover content within those domains except to introduce students to vocabulary and boundaries. We chose the business population as opposed to an engineering population to test specific elements in this study that have ramifications for engineering education, especially engineering entrepreneurship education. Similar introduction to first year courses in engineering at our university include significant content that focuses on teaching design skills and mathematical modeling.

Modeling, as noted in the use section, is a key engineering skill that may bias students towards predictive approaches. Conversely, the design thinking work might provide a push towards more effectual approaches because of close structural ties between the design and effectuation\textsuperscript{19}. This course content makes the engineering course interesting for future study, but inappropriate for
this study which focuses on the instruments’ development. Because the business introductory course does not address specific discipline content, but instead focuses on a broad overview of disciplines within management, it was seen as more appropriate for preparing the instrument for use on engineering students.

We used a pre-post study design. The first (pre) data collection occurred during the first class period, during the first week of the semester. For the first data collection, students completed the survey in class using paper forms, which were used to track class attendance of each student. The second data collection (post) occurred during the final week of the semester. Due to constraints in the class design, the second data collection occurred electronically and was gave students attendance credit for the final day of the course. In total, the two administrations of the test occurred approximately 15 weeks apart.

Overall, we received pre or post responses from 412 unique students. Within the 412 unique individuals, 337 completed the pre survey, while 381 responded to the post survey. In total, we had 306 complete responses (i.e., responses with a matching pre and post survey) and 106 incomplete responses (i.e., completed only the pre or the post survey). Due to fluctuations in the course enrollment the exact response rate is difficult to calculate accurately, although we estimate it to be approximately 90% due the overall class size.

In addition to collecting usernames to track completion of the survey, we collected two pieces of self-identified demographic information, gender and English language ability during the pre-survey. Of the 306 complete responses, 44% (134) self-identified as female while 56% (172) identified as male. We measured English language ability by asking the language of instruction of students’ most recent prior school. 91% of the students with both pre and post responses indicated that their most recent prior school taught in English.

**Analysis**

We used two approaches to analyze the collected data. The first approach was an examination of basic descriptive statistics to provide an over view and context of the distribution of student responses. The descriptive statistics used counts and means to assess how students responded to individual items and how responses changed across the semester in broad patterns. The descriptive statistics also allowed an overall assessment of data quality before proceeding with the second, more complex analysis stage.

In the second stage, we used a Rasch model to explore the structure and reliability of the items (Boone & Scantlebury, 2006). The Rasch model creates a probabilistic relationship between the ‘difficulty’ of a test item and the test takers’ ability (Bond & Fox, 2007). Because our instrument does not propose a correct answer but rather one of student choice, we adopted the term ‘effectuality’ instead of difficulty. Effectuality refers to whether, across the whole population or an individual, there was a shift towards more effectual answers.

The Rasch model approach is appropriate for our analysis for three reasons (Fraser, Johnson, & Templeton, 2011). First, the Rasch analysis makes use of an invariant basis, which tests the fit of data (both item and person) to a presupposed model rather than only providing a description of the underlying data. An invariant basis allows for a comparison of the individual students on the construct being assessed across the same test taken at different times. Second, it uses a logistic
regression rather than linear regression. Logistic regression is more appropriate than tools such as CFA that use linear regression because of the binary response model in the instrument which matches the inherently binomial response patterns of the binary item construction. Third, as opposed to IRT models which treat items separately, the Rasch approach looks at items as a collective unit. The collective approach is useful to assess the internal structure of the instrument and the coherence of items written for different heuristics on the overarching construct of effectual logic. It is also more extensible in later studies to a multi-scenario instrument.

A further benefit is that the Rasch analysis results include fit statistics, akin to confirmatory factor analysis. These fit statistics allow for an assessment of the consistency of item behavior (Wright, Linacre, Gustafson, & Martin-Lof, 1994) and the unidimensionality of FES, which is important because of studies showing that effectual measurement may or may not be a single dimensional construct. The fit values, called infit and outfit statistics, are evidence of how well our data matches the Rasch model’s expectations for variance and whether students’ actual responses fit the Rasch model’s probabilistic hypothesis about their performance (Bond & Fox, 2007). Both fit measures have a target value of one, denoting a solution with the exact amount of variance in responses expected by the Rasch model. Datasets with less randomness than expected smaller fit statistics, while results with more randomness result higher fit statistics. Acceptable mean square fit values for high stakes testing are 0.8 - 1.2 and 0.7 - 1.3 for lower stakes testing (Wright et al., 1994).

Results

Descriptive statistics
Prior to employing the Rasch model, we examined initial descriptive statistics within the responses. These included overall percent of effectual responses from students and the distribution and change over time of students’ responses. For the descriptive statistics, we chose to only use ‘complete’ responses that had both a pre and post survey response.

The percent effectual responses for each question in the both the pre and post data collection appear in Table 2. Notably, the items had a fairly wide range (26% to 81%) of effectual responses. However, the change in effectual responses changed very little (5% or less) between pre and post data collections.

<table>
<thead>
<tr>
<th>Item #</th>
<th>% Effectual Responses</th>
<th>Pre-post change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 4</td>
<td>81%</td>
<td>76%</td>
</tr>
<tr>
<td>Item 3</td>
<td>72%</td>
<td>73%</td>
</tr>
<tr>
<td>Item 5</td>
<td>51%</td>
<td>55%</td>
</tr>
<tr>
<td>Item 2</td>
<td>31%</td>
<td>33%</td>
</tr>
<tr>
<td>Item 1</td>
<td>28%</td>
<td>26%</td>
</tr>
</tbody>
</table>

Note: Items are sorted by the % effectual response on the pre survey.
We also looked at whether student responses stayed the same in the pre and post survey. The results, shown in Table 3, are useful for understanding the consistency of student responses. In comparing each individual person-item response, 63% of responses were consistent (i.e., the student chose the same response) between the pre and post survey. This percentage was considerably higher than the percentage of responses that changed from effectual to causal or causal to effectual. Additionally, the total number of effectual and causal responses stayed within 1% between the pre and post-tests.

Table 3 Cross tabulation of pre and post responses to items for completed responses.

<table>
<thead>
<tr>
<th></th>
<th>Post</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Causal</td>
<td>Effectual</td>
<td>Total</td>
</tr>
<tr>
<td>Pre</td>
<td>437</td>
<td>276</td>
<td>713</td>
</tr>
<tr>
<td></td>
<td>(29%)</td>
<td>(18%)</td>
<td>(47%)</td>
</tr>
<tr>
<td>Effectual</td>
<td>280</td>
<td>522</td>
<td>802</td>
</tr>
<tr>
<td></td>
<td>(18%)</td>
<td>(34%)</td>
<td>(53%)</td>
</tr>
<tr>
<td>Total</td>
<td>717</td>
<td>798</td>
<td>1515</td>
</tr>
<tr>
<td></td>
<td>(47%)</td>
<td>(53%)</td>
<td></td>
</tr>
</tbody>
</table>

We also looked at the distribution of ‘effectuality’, meaning the count of effectual responses by a given student on the pre and post-tests. The number of effectual responses ranged from 0 to 5 for each participant indicative of more causal or more effectual approaches to the scenario at the respective ends of the scale. A histogram summarizing the calculation appears in Figure 4. In general, the distribution is quite similar between the pre and post-test, which suggests that at the population level very little changed in general response patterns (i.e., the ‘effectuality’) of students across the semester. Because the students’ course work focused primarily on basic definitions of business concepts, and did not include entrepreneurial thinking, this is a result we would expect.

**Rasch Model Results**

The Rasch model fit results, shown in Table 4. The fit values in the pre-test ranged from 0.75 to 1.00 while values in the post test ranged from 0.82 to 0.93. The fit values are all within the range, 0.7 to 1.3) that are appropriate for typical testing. The values are all at or below 1.00, which indicates that there is less randomness in students’ responses than expected in the Rasch model.

The low fit values indicative of low variance in the data are not a problem in this case. They do not indicate increased measurement error, but rather a lower ability for the give set of questions to discriminate between similarly performing individuals. As we will discuss in the discussion, the lower discrimination capacity of individual items is of lower concern in FES than other instruments.

The fit values also indicate that the items function as a unidimensional construct. If an item did not function as a measures of a single construct, the infit and outfit values would be separated from the rest of the items. The fit values for all the items are generally consistent, and all fall into a similar range. Notably, the fit values that change from the pre to the post test generally moved
towards a better fit. The generally stable or improving measures of fit across the entire semester provides useful evidence of the reliability of FES measuring the intended construct.

Figure 4 Distribution of the number of effectual responses in pre and post tests

Table 4 Rasch model fit values for pre and post tests

<table>
<thead>
<tr>
<th>Item</th>
<th>Pre test</th>
<th>Post test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infit</td>
<td>Outfit</td>
</tr>
<tr>
<td>Item 1</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.83</td>
<td>0.75</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.90</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Beyond the fit values, the ‘difficulty’ values, which actually predict the likelihood of an effectual response, also indicate that the instrument fits the model fairly well. Items with an ‘effectuality’ of greater than 0 have a more than 50% chance of the average student selecting the effectual choice while items with an effectuality less than 0 have a lower than 50% chance of that student selecting the effectual response. The results, shown in Table 5, show that the items did not meaningfully shift in order or value of that parameter between the pre and the post test.
Only two items changed order within the effectuality results. Item 2 was the most likely to receive an effectual response while item 4 was the least likely for both the pre and post test. Item 5 was the only item which shifted its place in the order going from the 2nd most effectual item to the 3rd most effectual item. Item 5 is also the only item that switched from more likely to receive an effectual response (.12) to more likely to receive a causal response (-0.12) between the pre and post tests. Items 1 and 2 moved very little between the pre and post test (<0.05) while items 3, 4, and 5 moved somewhat more (~0.20) although this movement is likely attributable to the small number of items.

Table 5 Difficulty values of items in pre and post tests

| Item  | ‘Effectuality’ (in logit form) | | |
|---|---|---|---|---|
| Item  | Pre test (rank order*) | Post test (rank order*) | | |
| Item 1 | 0** (3) | 0** (4) | | |
| Item 2 | 0.87 (5) | 0.90 (5) | | |
| Item 3 | -0.80 (2) | -0.96 (2) | | |
| Item 4 | -1.38 (1) | -1.18 (1) | | |
| Item 5 | 0.12 (4) | -0.12 (3) | | |

* Rank order of difficulty is by test administration and goes from 1 (i.e., the item most likely to elicit a causal response) to 5 (i.e., the item most likely to elicit an effectual response) 
** the eRm package constrains the first item difficulty to 0 to allow estimation

Discussion and Conclusions

The purpose of this paper is to describe the methodology and preliminary steps required to develop and instrument referred to as the Factorial Effectuation Survey (FES). We are designing FES to measure students’ use of predictive or effectual decision making approaches when faced with differing business scenarios. To present a methodologically sound approach to instrument development, we sought to test three hypotheses in our pilot work: 1) the items in FES function as a coherent group to measure the intended trait, 2) all of the items in FES will provide useful information for separating students’ use of effectual and causal logic; and 3) students behavior on the FES measure will be stable across time at the individual and group level when students are not instructed in effectual, entrepreneurial, or business techniques.

In this paper, we created and executed an initial test to establish basic evidence for a portion of FES because the factorial survey model has multiple components that must operate in sync. We performed this test with business students in a class where the expectation was they would not change their decision making patterns to any great extent, but where concepts associated with the test would be introduced. We report our findings based on the three hypotheses below.

**Hypothesis 1: The items in FES function as a coherent group to measure the intended trait.**

Based on the fit values and the consistency of the of effectual and causal responses in the pre and post tests, we assess this hypothesis as true. The Rasch model results suggest that responses to the items are statistically predictable. That is a sign of coherent function of the individual items and the items as a group. Further, the fact that the number of student effectual and causal responses overall was quite consistent across the semester, despite changes in answers to
individual answers suggests that the items function cohesively to absorb variance in student responses.

We note that all of the fit values achieved were less than 1.00 (i.e., the expected amount of variance). This lower than ‘normal’ variance is a useful property in FES because the items will be taken multiple times by each respondent. Having lower variability in the item responses allows for the influence of the changes in context present in the full version of FES to be evaluated more precisely.

**Hypothesis 2: All of the items in FES will provide useful information for separating students’ use of effectual and causal logic**

As with hypothesis 1, we assess this hypothesis to be true. While the ‘effectuality’ of items is quite close, it is disparate and each item provides a different, and stable, measure of how effectual a student is.

We note that the information available from FES improved slightly (both in terms of precision of response targeting and fit) in the post test. This may suggest that the language of the items or the scenario may have been clearer to students later in the semester. This suggests that future version may need to carefully assess the level of knowledge required to understand and select among the responses within the scenarios that students are given.

**Hypothesis 3: Student behavior on the FES measure will be stable across time at the individual and group level when students are not instructed in effectual, entrepreneurial, or business techniques.**

The third hypothesis also appears to be generally true. Students in the sample were not given specific instruction on entrepreneurship or effectuation. Therefore, the lack of significant change across the semester is expected. Of the five items, no item moved more than 0.25 logit in ‘effectuality’ (i.e., the likelihood of an effectual response) from the pre to the post test. A 0.25 change is generally small and is within the typical precision in targeting item difficulty (akin to our effectuality) in Rasch modeling.

It is important to note that hypothesis 3 is focused not just on the precision of each item but of the measurement of students’ effectuality more broadly. The results comparing student behavior at the construct level changed little over 16 week semester, without instruction – which is a positive sign. The small changes that were observed can reasonably be explained by better understanding of vocabulary about business used in the instrument and scenario.

As was described in the intended use section, this work is not a full test of the FES instrument. However, given that the results show the items function in a reasonable way, we plan to progress with conducting further studies of validity evidence. Future work will expand the single scenario version of FES into a multi scenario version with all scenarios using the same items. This future work will assess how statistical measures, similar to those used in this paper, predict behavior and indicate fit of the items and scenarios to an intended structural model of the trait we are measuring, how students chose effectual or causal logic. In addition to the statistical tests, we plan for future work to include studying the readability and form of the items. While the items here show indication of function, we plan to revise them further using an expert panel study to gather evidence of the items alignments with the underlying effectuation construct definition. We
also plan to test the revised items with students in a qualitative form to ensure that they use terms and sentence construction that are clear to students representative of the test taking population.

In sum, the future work continues the process of constructing an argument for validity of FES’ use as an instrument to aid those teaching effectuation in their classrooms. They will give the initial evidence to support FES’ use but not a final one. Because the process of creating a validity argument is ongoing, we will continue to urge future users of FES to perform their own analyses to ascertain that the evidence collected here continues to be representative of how the items in FES function.

References


4. Krebbers, L. The first steps towards a quantitative measurement scale of Causation and Effectuation in a non-entrepreneurial student context. (University of Twente, 2015).


Appendix: Instrument

Note: This version of the instrument is not current. We advise those interested in using the instrument to contact us for the most current version of the instrument and instructions for its use.

For the five questions below, we would like you to pretend that you have decided to create a startup company today. Each question below has a partial statement and then two options to complete the sentence or phrase. Please select the option that you think best completes the sentence.

4) I would begin working on my startup by…
   - determining my goals, my plan to achieve them, and the resources I need to execute the plan.
   - considering what I know, what resources I have, and what I can do with those things.

5) When deciding on what initial steps to take I would focus on…
   - what steps will maximize my potential for future revenue or profit.
   - what steps will limit my potential losses in the initial stages.

6) When identifying early customers and partners I would…
   - focus on analyzing my competitors so that I can identify their weaknesses.
   - engage with people I trust and build partnerships even if they might later become rivals.

7) I would consider failed experiments or unintended outcomes to be…
   - indicators that the my companies decision making processes need improvement.
   - opportunities to learn about what my startup could do better.

8) At this early stage, it would be better for my company to pursue opportunities where...
   - I can predict whether our product will work, even if I cannot affect the outcome
   - I can affect the outcome of the products, even if I cannot predict that outcome in advance