How to Incubate Your Dragon: Methodologies for Collecting and Analyzing Data on Learning in Academic Makerspaces

ISAM 2019 Paper No.: 15

Morgan Chivers[1] and Martin K. Wallace[2]

[1] Morgan Chivers; University of Texas at Arlington Libraries; e-mail: morgan@uta.edu

[2] Martin K. Wallace; University of Texas at Arlington Libraries; e-mail: martin.wallace@uta.edu

Introduction

Hours after discussions at the 2nd International Symposium on Academic Makerspaces were set ablaze by Rosenbaum & Hartmann's clarion call for data dragons [1], we presented the origin story of the Maker Literacies initiative at UT Arlington [2] on our quest to gather and analyze relevant data to better understand the high-level, transferrable skills-building that goes hand in hand with making. As promised, we returned the following year to report on the scaling and honing work accomplished as part of an Institute for Museum and Library Services (IMLS) planning grant [3]. During this pilot program, professors and librarians at five universities across the United States tested and contributed to the refinement of assessment strategies for maker-based competencies that allow educators to provide evidence of learning in both affective and cognitive domains [4].

Our published work until now has not been about the data collected, per se. Data collection and analysis were not required for the IMLS planning grant; our primary goal for that phase of the project was to explore a variety of assessment tools and techniques for measuring student learning. The pre- and post-project self-assessment survey methodology we developed - the subject of this paper - is one such assessment technique we were able to test. Our data collection, processing, and analysis efforts have been in the interest of developing a robust proof of concept for measuring student learning in academic library makerspaces; as we embark on the open scholarly adventure of sharing *not only the results of our data analysis but the full dataset itself*, we hope this will nurture a communal research response to an identified need in the field for empirical studies of makerspace learning beyond Engineering [5], and even beyond STEAM.

Other research projects in the academic makerspace community are tending to fledgling data dragons of their own; we trust that we will collectively train our respective dragons to help humans share resources and insights in order to construct a more complete sense of the manifest learning we all facilitate in our daily lives in the lab. The physical and conceptual situating of the UTA FabLab in an academic library generally guides us through different terrain than our colleagues who are situated within academic departments when determining the scope of our research questions; nonetheless, ample opportunities abound for collaborative learning and comparison of methods and results along our journeys. Particularly notable precedents which will inform subsequent revisions of our methodology include the direct engagement with student perception of interdisciplinarity as found in Rosenbaum & Hartmann's ingenious 2018 recontextualization of Carberry et al.'s Engineering Design Self-Efficacy [6].

Our Maker Literacies beta-phase survey questions and data collected from Fall 2017 – Fall 2018 have been published to the Mavs Dataverse Data Repository [7] with the disclaimer that we do not believe it to be ideal or broadly generalizable at this stage, primarily due to the inherent fact that the foundation of this data set is our beta-phase list of maker competencies and surveys. As a function of the conclusion of the IMLS planning grant, the beta-phase maker competencies were formally rescinded and superseded by a revised list [8], informed by feedback from participants and internal work to refine our methodology while bringing it in line with existing best practices as identified in the literature. The original beta list of competencies is still available for reference [9], and would be particularly useful for those interested in exploring and analyzing pilot program data.

We believe we have successfully developed a sound, while certainly not perfect, methodology for measuring the learning taking place in academic library makerspaces using student self-assessment. As we continue to cultivate our methodology and collect more data, we are confident the data will become more reliable and generalizable, with the ultimate goal of establishing (inter)national standards for assessment of transdisciplinary, hands-on learning experiences. This paper details our data collection, processing, and initial analysis methodologies, as well as a frank discussion of criticisms of our approach, lessons learned, and a first look at trends emerging from that data. We conclude with a listing of use cases for the data that we are currently exploring.

Context

After a year of increasingly close coordination with faculty interested in assigning students to complete projects in our FabLab, UTA Libraries formed the Maker Literacies Task Force to develop a program for formal curricular coordination with faculty, focused on broadening our scope across disciplines and assessment of makerspace learning. When the first year of Maker Literacies pilot classes yielded feedback that was not usefully comparable between courses, we discerned the need for a standardized measurement tool to deploy across all courses participating in the Maker Literacies program. We decided upon a pre- and post-project self-assessment survey combination, including a range of question strategies such as multiple-choice options and open-ended responses. The majority of the survey structure is Likert scale questions asking students to rate their level of experience on various dimensions of the (then) eleven maker competencies. The Likert scales are the primary measurement apparatus considered for analysis herein; the open-ended questions did not reveal much generalizable insight, even with those students who attempted substantive responses.

We introduced the surveys informally in the Fall 2017 semester in six classes at UTA, without IRB approval or students' informed consent; instructors either required their students to take the surveys or offered participation credit to students who completed the surveys voluntarily. As part of the IMLS grant process, we sought and received IRB approval to administer the surveys; in order to combine earlier data with the data collected during the grant period, we had to retroactively obtain informed consent from students who completed surveys prior to IRB approval. We attempted to contact all 85 of them by email; 54 of those students replied with consent, three replied to opt out, and 28 did not respond. We have excluded all survey responses associated with opt-outs and non-responses from our public release of the data.

All surveys administered as part of the IMLS grant pilot program and continuing beyond its formal scope (Spring '18-Fall '18) were conducted with proper IRB approvals, either through UTA or the institution where the course was taught.

In total, 765 students completed the pre-project self-assessment surveys, including students who consented to participate in the study and those who opted out. Of these, 717 gave informed consent to participate in the study and received a corresponding post-project self-assessment survey. Students who did not provide consent did not receive a post- survey and we removed their responses to the pre- survey from the public data. We also removed students who did consent to participate in the study and completed the pre- survey, but who then did not complete a post- survey. Of the 717 students who consented to participate, 346 completed the post-project self-assessment survey. The data associated with these 346 students is considered to be a complete dataset for the pilot/beta of our ongoing work, as no additional data will be collected using the beta-phase surveys. We are treating this beta-data as a practice dataset as we work through the process to improve our methodology for the ongoing phase of the program. Of course, there may be potential for mapping some of the data from this set to the next set in cases where particular competencies remained relatively unchanged.

For reference, we have posted the final IRB-approved versions of our pre- and post-self-assessment question banks with the public data on the Mavs Dataverse, along with a data dictionary to help other researchers understand what each column heading means in the data set. Using the self-assessment question banks in tandem with the data dictionary provides the best understanding of the data collected. The data dictionary is not perfect, but should be suitable; there are no plans at this time to make this more user-friendly, as we will need to compile a whole new data dictionary for the updated surveys that we began using in Spring 2019.

Comparison of the Pre- & Post- Self-Assessment Likert Scales

Actual pre- and post-project self-assessment surveys included Likert scale questions for makerspace equipment knowledge plus beta-competencies 1-6 and 9. Working with the professors of each course, we made decisions about which competencies to assess based on compatibility with their intended Student Learning Outcomes, not a pre-determined agenda on our part that would have ensured we assessed all the beta competencies; we wrote the questions for each competency as they were needed. After the beta-phase, we ran Cronbach alpha analysis [10] on the Likert scales using the beta-data and determined that all scales were valid with the exception of the scale for competency 5, "employs effective knowledge management practices". We have corrected for this in the revised surveys, among many other improvements.

Key to our data collection and analysis strategy is the use of a pre- measurement and a dual post- measurement: "Reflection" and "Now". During the post- self-assessment, students are asked to rate their current perceived competency after completing the project, with a follow up question prompting them to think back to the beginning of the semester and re-rate their competency as they remember it before completing the makerspace project.

Justification for "Reflection" and "Now"

It may seem odd to ask students to reflect back and re-rate the competencies of their former selves prior to completing a makerspace project. We built this aspect into our surveys to avoid an exhibited problem inherent with simply asking students to rate themselves at the beginning and then again at the end of the course. Namely, students with some experience in a given competency tend to over-estimate their competence at the beginning of the semester, because they either do not fully understand what they are rating themselves on, and/or because they do not yet know what they do not know [11].

As is demonstrated by the data, many students tend to rate their competence the same on the post- self-assessment survey "Now" as they did on the pre- self-assessment survey,

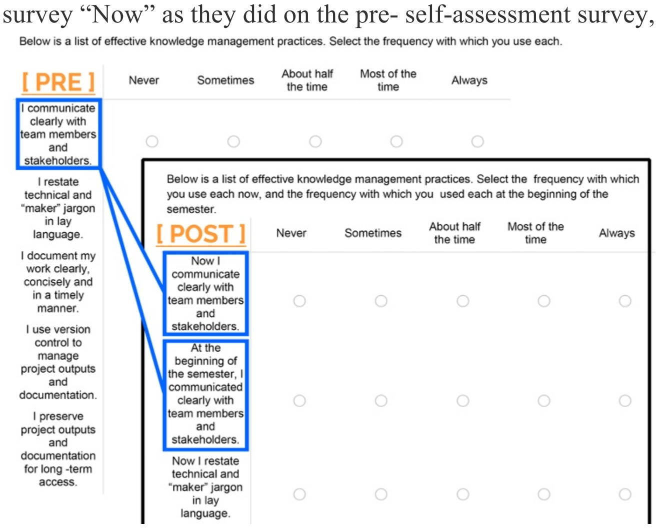

Fig. 1 Example of pre- & post- survey questions with Likert scales

which would indicate no growth in their perceived command of the skillset. In some cases, they even rate themselves lower on the "Now" than they did on the pre- self-assessment; if taken at face value, this might lead one to believe students are leaving the course with *less* competence than they had when they entered the course. By adding the "Reflection" measurements to the post-self-assessment survey, and comparing the "Reflection" to both the "Pre-" and the "Now" we are able to capture a more accurate understanding of students' perceived growth in competence between the beginning and end of the semester.

Well-noted criticism for asking students at the end of a semester to re-rate their perceived competence from the beginning of the semester is two-fold. First, the knowledge that students gained from completing the project will bias them when re-rating their initial competence, and second, students seeking social desirability will game the post-survey scales in order to show that they have gained competence because they believe that is what we want to see.

As for the first concern, we recognize that the bias objection is certainly pertinent when measuring objective knowledge, such as in pre-/post- testing where students are asked to provide correct answers to specific questions. We contend that this concern is not relevant when measuring subjective self-reported levels of competence. Indeed, the bias in the mind of the more experienced student regarding the relative state of their competence prior to the understanding gained through the course assignment is *precisely* what we are attempting to capture. Based on previous experience teaching, we were already aware that students tend to overrate their competence at the beginning of the semester and that we should not rely on the Pre- rating as an accurate measure; our data bears this out. We believe the "Reflect" and "Now" measures are more accurate than simply using a standard pre- and post-project self-assessments.

As for the second concern, we have revised the Maker Competencies surveys to incorporate Reynolds Short Form C of the Marlowe-Crowne Social Desirability Scale [12] in order to detect these types of responses and weight them accordingly in subsequent analysis. The data used in this paper's analysis is from the beta phase and thus has not incorporated any correction for social desirability bias.

Techniques for Validating and Analyzing Data

As stated, the beta-data is a testing ground for us to better understand data management and analysis techniques. Fundamental to the entire endeavor of Maker Literacies has been a full embrace of iterative design principles: the concept that one must make a best effort even in the knowledge that it will likely fail to achieve the envisioned potential on initial attempts, and that working through such inevitable issues is integral design praxis. This contextualization of failing intelligently as something to be expected, rather than an indicator of deficiency is the ether in which objects are created in response to project prompts, curriculum is revised between semesters, and the program itself evolves. This has been no less true for the processes of cleaning and analyzing our data.

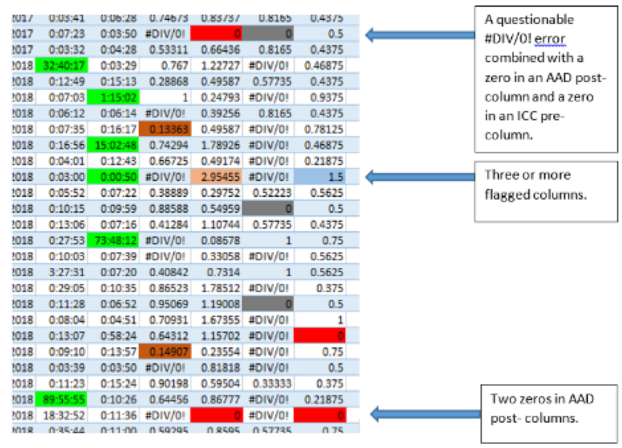

Fig. 1 Example of flagging and sorting out potentially unreliable data

It would seem to be self-evident folly to accept all gathered data as automatically reliable [13]; we have employed a variety of data validation techniques in order to better understand the nature of the dragon we are rearing. Analyzing data points to identify responses with characteristics of randomness or inattentiveness is essential to preserving the reliability of insights gleaned from analysis results. What is perhaps less clear is exactly how to best identify and remove unreliable results without bias or unsustainable efforts.

Here two roads diverged in a wood, and we took both. Statistical methods offer the allure of scalability, which is crucial as we prepare for significantly expanded streams of student learning data in current and future phases of the program. Math-based processes also tend to carry an air of objectivity, though we found that subjectivity was potentially a more pronounced factor in determining unreliability using the statistical methods we have explored thus far (more on this later). A manual approach offers a nuanced view of student responses and greater ability to discern reliability within an individual student's answers, but at the cost of long spans of tedious time that would be impractical for larger datasets. In this case, the manual coding also enables a significantly more profound view of the learning students experienced.

Statistical Heuristic for Identifying Unreliable Responses

One way of framing this is to look at the dataset with the understanding that we are comparing two sets of answers to the same sets of questions from the same student—once before the student completes a project in the makerspace, and once after the student completes their project. Generally, while one expects some variance, one would not expect wildly different answers in the second set, overall, from the first. Intraclass Correlation Coefficients (ICC) and Average Absolute Deviations (AAD) are ways to detect variance in a student's responses that might indicate unreliability.

Intraclass Correlation Coefficients show the strength of a relationship between two sets of data – in our case, comparing a set of answers to a series of pre-project self-assessment survey questions to a corresponding set of answers in a series of post-project self-assessment survey questions – and whether the relationship is direct or inverse. ICC is a probability ratio ranging between -1 and 1; the further the number is from zero, the stronger the correlation is, and the more reliable it is presumed to be. Note that ICC does not evaluate each individual answer in pre- compared to its post- counterpart. Analyzing for ICC involves grouping respondents within a given course to look for a pattern; isolating those that fall well outside this pattern for closer examination on a case-by-case basis.

Average Absolute Deviations, on the other hand, show a measure of the average variance between individual data points in two sets of data. In the case of AAD, the higher the number—the more variance—the more indicative of an unreliable response. AAD is not looking for a pattern, but at the actual variance in values among individual data points between two sets. In our case, within a set of answers to a group of questions from the pre-project self-assessment survey, we compare each answer directly to its corresponding answer from the post-project self-assessment "Reflection", and then average all the differences to calculate AAD.

Both ICC and AAD are probabilistic variables prone to subjectivity; the threshold for what gets flagged as "potentially unreliable" and what makes the cut is a matter of discretion deserving of further investigation beyond the scope of this paper. We have been generally conservative regarding the removal of data through these techniques.

A third metric initially used to identify potentially unreliable responses was survey duration; we used a histogram to identify notable outliers and flag those as potentially unreliable. A survey response that lasted less than one minute, for example, could implicate a student who quickly clicked through the survey without reading the questions. A survey response that lasted hours could imply that a student was trying to respond to the survey while distracted with other activities and not giving their full attention to the survey, though a long duration might also be caused by a student starting the survey and then deciding to complete it later when they could devote their full attention. Long survey durations are thus an inconclusive measure of reliability, while extremely short survey durations, especially when combined with other flags, might be sound reason to assume unreliability. All these considerations remain true for survey durations in general, though we later discovered an error in the way durations were converted from Qualtrics to Excel, and did not use survey durations as a factor in determining reliability for the purposes of this study.

In this paper, all examples are drawn from the data for IE 4340, an Engineering Project Management course that measures for knowledge of makerspace equipment, as well as competencies number 4 (assembles effective teams) and number 8 (demonstrates understanding of digital fabrication process). We chose IE 4340 as our example as it represents a robust initial case study from the beta-data: we ran virtually the same assignment in four sections across three semesters (Fall 2017, Spring and Fall 2018) with 52 consenting participants who completed both pre- and post- surveys.

We have prepared in-depth, step-by-step instructions as an appendix using the "knowledge of makerspace equipment" survey responses. In Appendix A, we explain how ICCs and AADs were calculated, the criteria used for flagging responses, and general guidelines for removing responses; this is intended to help illuminate our process, its value and inherent shortcomings, as well as encourage other researchers to get involved in this data analysis process.

Manual Coding of Response Sets

Another way of framing this is to look at the dataset as a series of triple responses. Every participating student provided a response to each measured dimension three times: Pre-, post- Reflect, post- Now. Whether one is screening for potentially unreliable responses or looking for meaningful measurement of the growth experienced as a result of the project, the continuity of an individual's responses across all dimensions, how their responses relate to those of their classmates, or even the overall competence the students reported, are potentially not as informative as a comparison of the three numbers in each triple response to each other by category of Pre- responses, Reflect responses, and Now responses. Each triple-set of Pre-/Reflect/Now responses was assessed independently for the perception of growth, non-growth, or regression, and assigned one of seven designations:

Table 1 Manual Coding Designations of Triple Response Sets

No Growth	all three responses identical
Standard Growth	Pre- & Reflect same/linear, growth in Now
Growth without Growth	difference between Pre- and Reflect, identical Reflect and Now
Apparent Regression	Now is the lowest score
Potentially Unreliable	Reflect is the highest score
Solid Growth	Pre- higher than Reflect, Now is highest
Aware of Ignorance	Pre- is higher than both Reflect and Now
Aware of Growth	Pre- and Now identical, growth between Reflect and Now

Tri-sets designated as <Apparent Regression> prompted a more holistic consideration of that student's entire response, including comparison across dimensions, and an evaluation of their short answer responses. Appendix B has been prepared to share a detailed look at this process.

Justification for Removing Potentially Unreliable Responses

Regardless of whether statistical or manual coding methods are used, a potential criticism warns to avoid removing any results from the data, lest we be accused of cherry picking data, or more seriously, *p-hacking* our data in order to achieve statistically significant results. We took the approach of trying to remove examples that were clearly the result of a survey not taken seriously, and left all other responses in the dataset.

The "Raw" data, the statistical method's "Reliable" data (as determined by subjective heuristic criteria), and the manual coding method's "Pruned" data are presented as part of an open question as to which of the three versions of the data is more accurate. For our present purposes, which mostly pertain to program improvement and demonstration of proof of concept, we believe that the "Reliable" data is indeed more reliable than the Raw, though our current preliminary analysis does not exhibit a statistical significance between the two statistically sorted versions of the data. An understandable criticism of data validation methodology that does not yield statistically significant results when compared with the Raw data is that one might just as well save the effort of performing

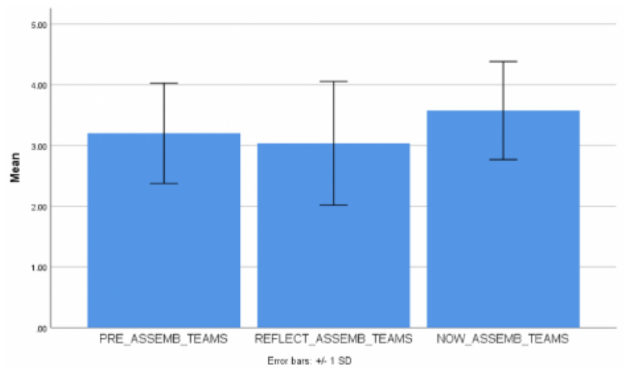

Fig. 3 Raw beta-Competency 4 data from IE 4340, aggregated

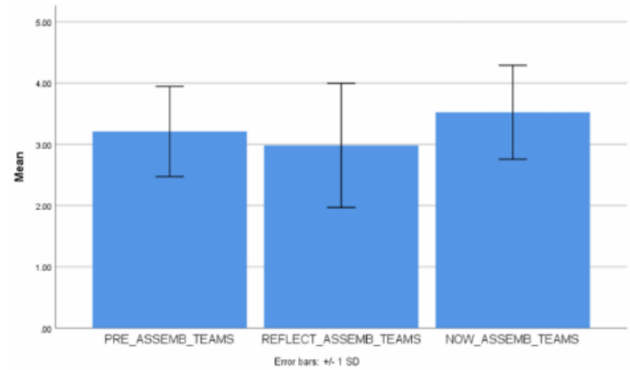

Fig. 5 "Reliable" beta-Competency 4 data from IE 4340, aggregated

the AAD and/or ICC calculations if the Raw results yield effectively the same insights. While that is true for this particular course we used as the case study from the overall beta-dataset, we anticipate other courses and larger datasets would not necessarily hold this pattern.

The "Pruned" data shows slightly more differentiation between the three response types when averaged, though still not a sufficiently significant difference between the raw data and the "Pruned" view to warrant the extra labor, *if* removing potentially unreliable responses was the only valuable function of manual coding; the manual coding method proves far more worthwhile in the data analysis process.

Other Considerations

In addition to the above problems and noted criticism, we have either already reconfigured our surveys to take into account the following circumstances, or we are actively seeking solutions. Each of these has been identified as problematic by the literature; we are implementing recommended best practices from that same literature in order to improve our survey methods. Most prominently:

- "Likert scale left bias" indicating that people are more likely to select the response options located on the left side of the scale [14];
- "Negatively worded stems" indicating that negatively worded stems tend to confuse individuals, causing them to select the responses that are opposite to their beliefs [15];
- and "Bidirectional response options" intended to help identify potentially unreliable responses [16].

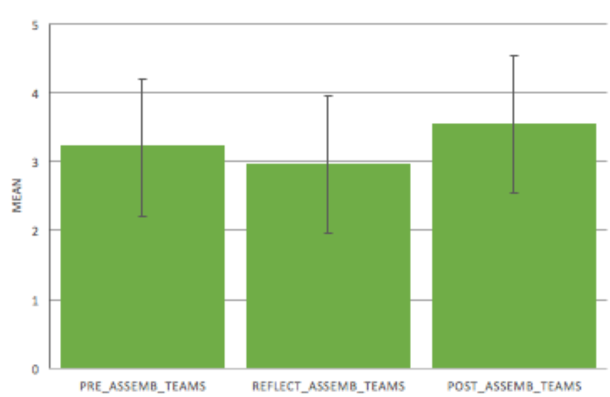

Fig. 4 "Pruned" beta-Competency 4 data from IE 4340, aggregated

Statistical Data Analysis

Taking the above criticisms into consideration, we present six statistical "views" of data from IE 4340 through the lens of compared averages. In our surveys, each of the selected dimensions are presented as Likert scales where students rate their competence on a scale of 1-5, where 1=no competence and 5=highly competent. For the comparison of means, we calculated a student's score by averaging the four dimensions together, and we calculate aggregate scores by averaging all student scores together. These comparisons use aggregate scores to depict six views in three bar graphs.

A. Six Views of the Data

- Raw Pre- and Post-/Now
- "Reliable" Pre- and Post-/Now
- "Pruned" Pre- and Post-/Now
- Raw Post-/Reflection and Post-/Now
- "Reliable" Post/-Reflection and Post-/Now
- "Pruned" Post/-Reflection and Post-/Now

B. RAW Assembles Effective Teams

- N=52 students
- Pre: 3.202 Post/Reflection: 3.039 Post/Now: 3.577

From Fig. 3 we see that, on average, students overestimated their competence in assembling effective teams in the pre-project self-assessment survey by a percentage difference of 5.224% (comparing Pre- to Reflect), and gained a percentage change of 17.703% in their perceived ability to assemble effective teams (comparing Reflect to Now).

C. "RELIABLE" Assembles Effective Teams

- N=43 students (nine removed unreliable responses)
- Pre: 3.209 Post/Reflection: 2.983 Post/Now: 3.523

From Fig. 5 we see that, on average, students overestimated their competence in assembling effective teams in the pre-project self-assessment survey by a percentage difference of 7.3% (comparing Pre- to Reflect), and gained a percent change of 18.103% increase in their perceived ability to assemble effective teams (comparing Reflect to Now).

D. "PRUNED" Assembles Effective Teams

- N=48 students (four removed unreliable responses)
- Pre: 3.208 Post/Reflection: 2.953 Post/Now: 3.542

From Fig. 4 we see that, on average, students overestimated their competence in assembling effective teams in the pre-project self-assessment survey by a percentage difference of 8.278% (comparing Pre- to Reflect), and gained a percentage change of 19.946% increase in their perceived ability to assemble effective teams (comparing Reflect to Now).

Note that the differences in this case study between Raw "Reliable" and "Pruned" data are not statistically significant.

Manual Coding Data Analysis

While the statistical approach relied on mean calculations to provide broad insights into the learning that took place within an entire class, it is a relatively blunt instrument for analyzing this sort of assessment data. By comparison, the manual coding process yields significantly more detailed views of the learning that took place within each student's experience of every assessed dimension separately. More to the point, manual coding enables us to orient our investigation towards perceived sense of growth within the student, rather than an assumption that individuals within a class are effectively reducible to the cumulative experience of the class.

If we are oriented towards measuring growth and not overall competence, then averaging all responses obscures our ability to view growth by focusing on the total recorded result as compared to the maximum potential rating. By coding each triple-response with markers that are only concerned with the perception of growth, an entirely different vista takes shape.

The four sections of IE 4340 were all asked to provide Pre-Reflect and Now responses to eight makerspace equipment dimensions and four team building competency dimensions. This results in 383 tri-sets of equipment knowledge and 192 tri-sets within beta-competency 4 (team building), after removing the Pruned unreliable results.

If we had performed a standard Pre-/Post- survey – assuming those responses would align to our Pre- and Now data – we would have seen that 36% of participating students indicated growth with makerspace equipment, and only 24% of students perceived growth in their team building skills. Furthermore, it would seem that almost 15% of these students *lost* skills in makerspace equipment, and almost 12% felt they *lost* competence in their team-building abilities.

These would not be encouraging results for any pedagogy.

The added reflective component of our survey methodology, however, when combined with the manual coding of individual tri-sets, unlocks a window into the student perception of growth that attempts to capture the experiential bias gained through the semester project, and correct for the Dunning-Kruger effect as present in the Pre- responses.

Through this focused lens, we see instead that over 41% of students perceived growth in the makerspace equipment skills, and over 47% of students perceived growth in their ability to build effective teams.

From a pedagogical perspective, these numbers of perceived growth are lower than we would hope for, and point to an opportunity for continued iteration of this project prompt to more fully engage every student. From an assessment perspective, the 23% shift in measureable perception of student growth between standard Pre-/Post- analysis and this innovative Pre-/post-Reflect/post-Now survey methodology is encouragingly validated by this analysis.

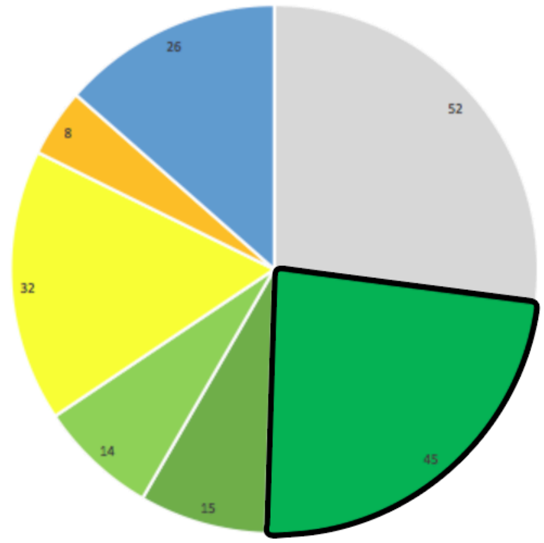

Fig. 6 Pruned beta-Competency 4 data from IE 4340, tri-set analysis (please refer to Table 1 and Appendix B for key and additional context)

Conclusion

We started this project with three hypotheses:

Hypothesis A – students tend to overestimate their competence when completing the pre-self-assessment survey – holds true, as exemplified in the raw, "reliable", and pruned views of the data.

Hypothesis B – the Reflect methodology within the post-project self-assessment surveys will yield additional insight regarding student self-perceptions of growth that would have otherwise been recorded as non-growth in a typical pre-/post-survey – is strongly supported by this analysis. Self-perception is an ever-moving target; by building in reflective components to the surveys, students are prompted to think about the differences they see in themselves across the span of a semester and provide responses that reflect this growth.

Hypothesis C – that students who do not take surveys seriously will provide bogus data that will skew our results if not fastidiously removed – is not supported by the analysis of this course's data. Surprisingly, Raw, "Reliable", and "Pruned" data all show similar patterns with negligible differences after filtering for potential unreliability. We anticipate seeing a more significant difference between raw and filtered data going forward, with larger datasets and more nuanced methods for filtering and analyzing responses.

In conclusion, based on this analysis, we can substantiate that assessing for the triple-set of Pre-/post-Reflect/post-Now is a more accurate data methodology for measuring student learning than using Pre/Post questions alone, and that our methodologies for sorting raw data into reliable data remain an open question in need of further review.

References

[1] L.F.Rosenbaum & B.Hartmann, "Where be Dragons? Mapping the known (and not so known) areas of research on academic makerspaces," *Proceedings of the 2nd International Symposium on Academic Makerspaces*, Cleveland, OH: Case Western Reserve University, 2017. https://drive.google.com/open?id=0B4ZIatyugWjJY1ZjOXhIVGRzSFU

[2] M.K.Wallace, G.Trkay, M.Chivers, & K.M.Peery, "Making Maker Literacies: Integrating academic library makerspaces into the undergraduate curriculum," *Proceedings of the 2nd International Symposium on Academic Makerspaces*, Cleveland, OH: Case Western Reserve University, 2017. http://hdl.handle.net/10106/27017

[3] M.K. Wallace, G. Trkay, K.M. Peery, M. Chivers & T. Radniecki. *Maker competencies and the undergraduate curriculum* (IMLS #LG-97-17-0010-17), 2017. https://www.imls.gov/grants/awarded/lg-97-17-0010-17

[4] M.K.Wallace, G.Trkay, K.M.Peery, M.Chivers, & T. Radniecki, "Maker Competencies and the Undergraduate Curriculum," *Proceedings of the 3rd International Symposium on Academic Makerspaces*, Palo Alto, CA: Stanford University, 2018. http://hdl.handle.net/10106/27518

[5] L.F.Rosenbaum & B.Hartmann, "Making Connections: Project courses improve design self-efficacy and interdisciplinary awareness," *Proceedings of the 3rd International Symposium on Academic Makerspaces*, Palo Alto, CA: Stanford University, 2018. https://drive.google.com/file/d/1CxlaiVFZqTiUCBDBHt_wGWiMvT4ry-WV/view

[6] A.R.Carberry, H.S.Lee, & M.W.Ohland, "Measuring Engineering Design Self-Efficacy," *Journal of Engineering Education*, vol.99, 71-79, 2010.

https://ceeo.tufts.edu/documents/journal/carberry_lee_ohland.pdf

[7] M.K.Wallace, "Maker Literacies Student Learning Data," *Texas Dataverse Data Repository*, 2018. https://doi.org/10.18738/T8/ZCZF6X

[8] M.K.Wallace, G.Trkay, K.M.Peery, M.Chivers, & T. Radniecki, *List of Maker Competencies, Including Preamble and Acknowledgments*, 2018. http://hdl.handle.net/10106/27634

[9] UTA Libraries Maker Literacies Task Force, *Beta-List of Maker Competencies*,2016. https://library.uta.edu/sites/default/files/MakerCompetenciesList.pdf

[10] L.J.Cronbach, "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, vol.16, 297-334, 1951. https://doi.org/10.1007/BF02310555

[11] J.Kruger & D. Dunning, "Unskilled and Unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments," *Journal of Personality and Social Psychology*, vol.77, no.6, 1121-1134, 1999. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0022-3514.77.6.1121

[12] W.M.Reynolds, "Development of reliable and valid short forms of the Marlowe-Crowne social desirability scale," *Clinical Psychology*, vol.38, no.1, 119-125, 1982. https://doi.org/10.2466/pr0.1988.63.1.227

[13] A.D.Horowitz & T.F.Golob, "Survey data reliability effects on results of consumer preference analyses," *Advances in Consumer Research,* vol.6, 532-538, 1979. http://acrwebsite.org/volumes/9612/volumes/v06/NA-06

[14] A.Bandura, "Guide for constructing self-efficacy scales," *Self-Efficacy Beliefs of Adolescents*, vol.5, no.1, 307-337, 2006. https://www.uky.edu/~eushe2/Bandura/BanduraGuide2006.pdf

[15] J.J.Barnette, "Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems," *Educational and Psychological Measurement*, vol.60, no.3, 361-370, 2000. https://doi.org/10.1177/00131640021970592

[16] H.Maeda, "Response option configuration of online administered Likert scales," International Journal of Social Research Methodology, vol.18, no.1, 15–26, 2015. http://dx.doi.org/10.1080/13645579.2014.8851

How to Incubate Your Dragon:

ISAM 2019 Paper No.: 15

Appendix A

Methodologies for calculating ICC and AAD in Excel

This guide is provided to help illuminate our process for identifying potentially unreliable responses in the Maker Literacies Public Pre/Post Combined Assessment Data for Fall 2017 - Fall 2018[i]. To aid readers in understanding our methodology, a sorted version of the data in a revised Excel file[ii] is included, in which three criteria have been employed to help identify potentially unreliable results: intraclass correlation coefficient (ICC), average absolute deviation (AAD), and survey response durations from both pre- and post-self-evaluations. Take a moment to download both data files and open them to follow along.

Note that the revised version of the data will not be placed in the Texas Data Repository Dataverse. The version shared here is for example only; other researchers are encouraged to take the original data and use their own criterion thresholds to remove potentially unreliable responses. Since the criteria for determining which responses might be unreliable are subjective, others may want to repeat the process to check our conclusions, or tweak their thresholds to be more restrictive than this conservative example.

Scrolling down the left side of the Data sheet, notice that many cells are highlighted with an assortment of different colors. Some rows have one or two colors; some have more dotting the cells further to their right. In total, there are eighteen different flag colors: eight ICCs (equipment knowledge, competencies 1-6, and competency 9), eight AADs (equipment knowledge, competencies 1-6, and competency 9), and two durations (pre-survey and post-survey). Each color is a flag indicating that the response is among the least reliable responses according to one of the three criteria: ICC, AAD or duration. The more colors tagged in a row, the less reliable that respondent data is among all potentially unreliable responses. Rows that have three or more color tags are candidates for removing from the data outright. Rows that have fewer colors may have unreliable data in the flagged ICC, AAD, or duration sub-set, but still have reliable data in its non-flagged areas. In this sample data, all rows have been left intact with none hidden or deleted. Reliability of each row will be determined on a case by case basis while performing the data analysis for each specific sub-set of data (i.e., for each competency). Having them flagged will help ease and expedite the analysis process.

Also note just how many rows are flagged with at least one color. If we were to remove every row that was flagged, it would only leave about a third of the total responses, and this after using very conservative measures for placing responses into the potentially unreliable category (more on this below).

The process for calculating ICCs and AADs required first copying the target data sub-set from the master Data sheet into two separate sheets, one for ICC and one for AAD (those sheets are intact as tabs in the example), deleting and rearranging some columns, hiding empty rows, and then using some of Excel's built-in functions on the remaining data to get a discrete set of data for every competency measured into its own Excel sheet. Once the data has been copied and trimmed into new sheets, perform the functions for ICC and AAD, and copy the data calculated in those columns into new columns on the master Data sheet. Note: add eight columns for AAD; empty columns for all eight ICCs are already in the public data set.

Because this process is involved and relatively time-consuming, only the first iteration of calculating the ICC for one sub-set of data, the Knowledge of Makerspace Equipment set, will be described in depth. This identical procedure has been performed on all remaining sub-sets of data, including competencies 1-6 and 9.

Similarly, since the process for flagging potentially problematic AADs is extremely similar to flagging potentially problematic ICCs, only those aspects that are different from ICC will be explained in depth, while referring to relevant steps from the ICC examples when possible.

In-Depth Instructions

First, download and save the data as an Excel file, and then create a copy of the file by saving it "as" with a new name to preserve the ability to easily return to the original file to compare results or begin again.

A. Intraclass Correlation Coefficients

For "Knowledge of Makerspace Equipment" and each Competency (1-6 & 9), calculate an ICC following this process; the provided example uses the makerspace equipment data:

1. Freeze the pane to right of target ICC column so that target ICC column remains static when scrolling left and right. If it is already frozen at another location, first unfreeze the pane, then try again.
2. Scroll to the right to the first column of the target data. In the example provided, that is the heading "Pre.knowledge of makerspace

equipment - 3D Printing". Select this and all of the target Likert scale data headings to its right. In the example, we want all Likert scale headings about knowledge of makerspace equipment selected. These will always be consecutive columns that can be selected using one Shift-click to grab them all.

- Do NOT select the multiple choice or text entry data that may be before or after the Likert scale data. In all cases, Likert scale data will be groups of 3 columns: a pre- column, and post-before/beginning column, and a post-now/end column.

3. With the column headers still selected, scroll to the very bottom of the data, row 347. Hold down Shift and click in the cell in the bottom row (row 347) directly beneath the last column heading in the target data. With all target data selected, use Ctrl+C to copy the data.
4. Open a new sheet/tab and rename it to match the ICC of the current target data, for example, Equip_ICC. While in cell A1 in this new sheet, use Shift+F10 + V to paste the values (not the equations) that were copied from the Data sheet. Do not unselect the data; immediately go to the Insert tab and select Table. If Table is not available (if it's "grayed out") then the data miraculously pasted as a table.
    - When inserting the table, be sure all pasted data is still selected. It should remain selected after pasting, but if the data was unselected, go back and select it again.
    - Since this process also copied/pasted the column headings, be sure to select the checkbox "My table has headers".

For conducting ICC and AAD analysis of two data sets, we want to use the pre-self-assessment as one data set, and specifically *the reflective versions* of the same questions from the post-self-assessment as our second data set. Therefore:

5. Delete all post-now/end columns from the data (every third column), leaving only data from the pre-survey and post- before/beginning columns.
6. Separate pre- columns from post- columns into two groups, while keeping everything in order. There should be an equal number of pre-columns on the left, and post- columns on the right.

The next step is to flag and hide all the rows where either the pre- or post- (or both) data is completely empty, though configuring the work environment for the remainder of work ahead will make the remainder of this tutorial significantly easier, avoiding the trap of scrolling left and right endlessly.

7. Select the entire table using the southeast pointing arrowhead in the top left of the spreadsheet. With the whole table selected, click and drag column edges until all columns appear within view. Since all columns are selected, shrinking the size of one column, will automatically shrink the other columns to the same size.

With the columns compressed, it is much simpler to scroll down the sheet to find empty rows, and rows where either the pre- or post- data is empty. These rows should be flagged as potentially unreliable.

8. Go down all 347 rows, flag the rows with a color of your choice, and hide them (select flagged rows > Format > Rows > Hide)
9. Now that all the data is in its proper place, insert a new empty column at the far left of the sheet, labeled ICC.
10. Select the top cell in the new ICC column. Paste the formula =ABS(CORREL(array1,array2)) into the formula bar.
    - Still in the formula bar, click on the word array1. A small helper bar will appear underneath with the formula. In that helper bar, again click on "array 1". Now click and drag across all the target pre- data in that row (B2 through W2 in the example).
    - Click again in the helper bar, this time on the word "array2". Click and drag across all the target post- data in that row (X2 through AS2).
    - Now both data arrays are selected. Let go of the mouse button. Don't click anything.

After doing the above, this is the example formula for the equipment ICC: =ABS(CORREL(B2:W2,X2:AS2)). As the example sheet is set up, data will always begin in cell B2; the rest will vary depending on how many columns are in the competency Likert data, but the structure of the equation should look similar.

11. Now hit enter. The Intraclass Correlation Coefficient should now appear in cell A2. Note that sometimes Excel will auto-fill the cells below with their ICCs, other times it must be done manually.
    - If it didn't populate automatically, fill the column down with the formula by selecting cell A2, then use

- Shift+Ctrl+down arrow all at once. This will select all the cells in the column. Then Edit > Fill > Down.

Since ICC is looking for distance from zero to determine a threshold for what is and what isn't considered reliable, append the ABS() function to get the absolute value of the ICC, to convert negative values to positive ones. This will allow better visualization of distances from zero in a histogram.

Do not be alarmed that many of the cells are populated with divide by zero errors (#DIV/0!). This error indicates that due to either the pre-answers or the post-answers having an average standard deviation of zero, the coefficient can't be calculated. This happens when all values in either the pre- or post- data are equal. Since nothing can be assumed by this data, we did not flag these as potentially unreliable. To avoid impacting further analysis of the correlation coefficients, however, hide all rows with this error.

12. Use the filter in the column header to uncheck the value #DIV/0!, leaving only numerical values in the data set.
13. Create a histogram of the remaining visible data in the ICC column to better visualize the potentially unreliable responses.
    - Select all the data (not the headers) in the column, switch to the Insert tab in Excel and select Histogram (or Bar) in the Charts pane.

For ICCs, the number of responses in the furthest-left column in the histogram holds the data rows with the lowest ICCs, representing the least reliable results. In the example, this results in 19 rows with values of 0.17 or lower.

14. Flag the rows with ICCs below the desired threshold by adding background color other than bright red, since that color is already being used to represent missing data. To do this without sorting the data again (which will result in mis-aligned data when adding flagged rows back the master data sheet), use the filter in the ICC column once again, and deselect all values greater than 0.17.

Setting the threshold is where subjectivity begins to play a part in this data processing. We have chosen to be very conservative by identifying only the bin to the far left as representing potentially unreliable results. Since we are dealing with probabilities, it is a potential that some of the data in the left bar are perfectly reliable and that some data in bins to its right contain unreliable results. It is important to keep in mind that the purpose of this process is to create a heuristic approach that will save time from looking at every response individually and subjectively determining the reliability of each response. One may decide to flag responses in the second, third, or even fourth bin; but remember, the more rows you flag, the fewer rows will remain in the sorted data and even the conservative measures outlined here already flag about two-thirds of all rows.

Before moving on, unhide all of the rows in the ICC sheet that have been hidden throughout this process—the empty rows, the rows where the respondent didn't answer the questions in either the pre- or post- survey, and the rows with the divide by zero errors. Skipping this critical step, will result in misaligned data when reintegrating with the master Data sheet.

15. To unhide all the rows, first use the filter in the column header and check Select All so that the values #DIV/0! and values greater than 0.17 once again appear, then select the entire sheet, right click anywhere in the data, and select Unhide to reveal the rows with missing data.
16. Copy the data (not the header) in the new column, then go back to the beginning Data tab.
    - Select the top empty cell in the target ICC column (not the header). Paste both the value (not the formula) and the background color from the ICC sheet, using Shift+F10. In the menu that pops up, hover over "Paste Special" and select "Values & Source Formatting". This will paste the values and color formatting from the ICC sheet.

This concludes the instructions for calculating ICC's for each student respondents' familiarity with makerspace equipment pre- and post-self-evaluations. Repeat this process for each competency in the data (competencies 1-6, plus 9) until all of blank ICC columns are filled in.

B. AVERAGE ABSOLUTE DEVIATION:

Since this procedure is nearly identical as the process for calculating ICC, only those steps that deviate from the instructions above are listed below.

1. Blank ICC columns were already available in the example worksheet. For AAD, add a new empty column to the right of each of the ICC columns. Label the column headers similarly to how ICC columns were labeled.
2. Open a new sheet/tab and rename it to match the AAD you are working with, for example, Equip_AAD. Follow the remainder of instructions in number 4, above, substituting AAD for ICC.
3. Select the top cell in the new AAD column. Paste the formula =AVEDEV(array1,array2) into the formula bar.
    - Still in the formula bar, click on the word array1. A small helper bar will appear underneath with the formula.

In that helper bar, click on "number 1". Now click and drag across all the target Pre data in that row (B2 through W2 in the example).

- Click again in the helper bar, this time on "number 2". Click and drag across all the target Post data in that row (X2 through AS2).

While there will be no divide by zero errors (#DIV/0!) when calculating AAD, there will likely be many zero (0) values. This means that student gave the same answer to every question, indicating that they simply clicked down the column in the Likert scales. This may be indicative of a student who didn't read the questions and rushed through the survey. We have chosen to flag these as potentially unreliable, but since this is subjective, others may opt to not do this.

4. Once flagged, hide the rows using the column header filter.
5. Create a histogram of the remaining visible data in the AAD column to aid in identifying remaining potentially unreliable responses, using the same process as for ICCs, above.

For AADs, the number of responses in the furthest-right column in the histogram holds the data rows with the highest AADs, representing the least reliable results (opposite of ICCs). In the example, this results in only 1 row with value of 2.51 or greater.

6. Flag the row in the data by adding background color other than bright red or the color chosen for the ICC column. Again, to avoid re-sorting and therefore mis-aligning data with the master data sheet, use the filter in the AAD column, deselecting all values lower than 2.51.
7. Unhide all the rows as described with ICC, above, then copy and paste the values with format back into the corresponding column in the master data sheet.

This concludes the instructions for calculating AADs for each student respondents' familiarity with makerspace equipment pre- and post-self-evaluations. Repeat this process for each competency in the data (competencies 1-6, plus 9) until all of blank AAD columns are filled in.

Now that potentially unreliable responses in the data have been identified and sorted out, the data is ready for initial analysis.

Future Work

We intend to dig deeper into this pilot program data even as we carry out the next phase(s) of the program. Using the answers to multiple-choice questions, known data (i.e. course number, discipline, etc.), and data from other systems (i.e. GPAs and other information stored in student records), we will subdivide our Likert scale findings further, providing insight into these and other queries:

- Compare multiple courses on a specific competency
- Compare multiple disciplines
- Compare students who have previously used the makerspace to those who have not
- Compare by demographic data (year of study, age, gender, etc.)
- Identify project prompts with greatest impact
- Determine if there is a correlation between student self-perception of competence and grades
- Construct a deterministic/predictive baseline index per category (competency, course, discipline, assignment/project prompt, etc.)
- Predict students who are most likely to identify as "makers" based on the above analysis, and take steps to further encourage them
- Predict students who are most likely to not identify as "makers" and attempt to engage them in ways that will spark their interest in iterative design

As you explore the dataset for yourself, we anticipate additional research questions beyond what we have outlined here. We earnestly invite other researchers to engage with the analysis of this dataset and to share findings as openly as we have here.

[i] M.K.Wallace, "Maker Literacies Student Learning Data," Texas Data Repository Dataverse, V2, UNF:6:7+VEu6f/pln3j0+vvQsILA==[fileUNF] doi:10.18738/T8/ZCZF6X

[ii] https://uta.app.box.com/s/ett6snwzoemkpjyrubseyxfr207g2357

How to Incubate Your Dragon:

ISAM
2019
Paper No.:
15

Appendix B

Methodology for Manual Coding of Triple-Response Sets

This guide is provided to help illuminate our process for identifying potentially unreliable responses in the Maker Literacies Public Pre/Post Combined Assessment Data for Fall 2017 - Fall 2018. This manual coding method provides another way to look at and work with the same dataset that was explored statistically in Appendix A.

Setting the Stage

The first requisite step is to format the data downloaded from the Dataverse to be more human readable. Simple procedures such as standardizing column widths and row heights, adding blank rows between courses, centering numbers within columns, etc, can go a long way towards making the experience far less of a strain on one's eyes and mind.

As stated in the paper, the provided file is a case study of analyzing the data from all the sections of IE 4340 taught over the span of three semesters, though the same processes would hold true for any class or classes one might choose to explore. The desired range of cells was selected for each block of IE 4340 data and copied into a fresh spreadsheet; one could also choose to hide all the undesired cells. Within the IE 4340-only data, hide or delete all the blank columns associated with survey questions not posed to this class. Note that best practices in this case would encourage hiding columns to preserve future comparability of column labels to results from other courses.

With the view condensed to only relevant fields, take a few moments to visually delineate between the groupings of three responses that comprise each set of Likert data for a given equipment or competency by adding column borders between the tri-sets.

Color Field Key

Table 1 Manual Coding Designations of Triple Response Sets

No Growth	all three responses identical
Standard Growth	Pre- and Reflect same/linear, growth in Now
Growth without Growth	difference between Pre- and Reflect, identical Reflect and Now
Apparent Regression	Now is the lowest score
Potentially Unreliable	Reflect is the highest score
Solid Growth	Pre- higher than Reflect, Now is highest
Aware of Ignorance	Pre- is higher than both Reflect and Now
Aware of Growth	Pre- and Now identical, growth between Reflect and Now

The process is relatively simple: go through each tri-set one by one and assign each group a color code based on the key listed above. For a relative measure of speed, code the more recognizable tri-set patterns with the same fill color selected before turning to the inevitable set-by-set evaluation, determination, and marking.

Additional context about the special delineations:

<Growth without Growth> This category encompasses response sets that would seem to measure growth between the Pre- and the Now responses, though the Reflect and Now responses are identical. This indicates the student did not actually perceive any growth when reflecting on the semester's efforts within that dimension. Students cannot see their responses to the Pre- when filling out the Post- survey; this response type was unanticipated. It will be interesting to code other courses within the beta-data to get a better sense of the frequency of this type of a response and attempt to understand the motivations behind this perspective.

<Aware of Ignorance> For the purposes of this paper, this category was not included in the reported figures regarding growth, though more consideration about the appropriate reporting conclusion seems warranted. Despite recording a self-perception of static skills in the dimension throughout the semester, the student's understanding of the concept evidently grew sufficiently for them to recognize that they were less skilled in that area than they previously thought; this is a potential indication of mental growth.

<Aware of Growth> This is the ideal result to reinforce the confirmation of Hypothesis B. The student who registers growth between the Reflect and Now, but who rates their competence in the Now as identical to the Pre, would seem to have a perception not only of their own increased skills, but also an at least latent awareness of the depth of familiarity required to attain proficiency in the measured dimension. Again, respondents are not able to see their Pre- responses when determining their Post- responses.

<Apparent Regressoin> These are problematic ratings as there are not many logical reasons one would report learning in this way. All tri-sets coded with this marker should initiate an examination of other tri-sets and short answers in the same row to attempt a determination of whether the student was paying attention to the survey or not.

<Potentially Unreliable> Try to avoid coding any individual tri-set using this delineation. Mark tri-sets by the six other codes, then evaluate those rows with <Apparent Regression> and perhaps <Growth without Growth> or other suspicious patterns for potential determination as to whether the entire row should be considered a <Potentially Unreliable> student response.

Specific Examples

All references line numbers within this Appendix refer to the spreadsheets below.

An example of <Apparent Regression> that was ultimately determined to be <Potentially Unreliable> is Line 11: eight of the total eleven tri-sets in the Makerspace Equipment category were <Apparent Regression>; all responses in both Post- responses for Competency 4 indicated <No Growth> except for one, which yielded an <Apparent Regression>; several of the short answer responses were decidedly unsubstantive.

Not all student responses with multiple <Apparent Regression> designations were deemed <Potentially Unreliable>. Line 31 provides a good example: two of the eleven Makerspace Equipment responses were <Apparent Regression>; responses for Competency 4 were believable; the short answer responses indicated a student who entered the course with ignorance "In my opinion, a 3D printer is an object that uses lasers to actually make a final prototype object at the end/ I don't have any experience with CNC technologies work, I honestly do not know," and left with a reasonable, if novice, understanding of the principles "A 3D printer is an instrument that converts a digital file to a prototype/ [CNC] Is a computer that converts the design into numbers."

In the manual coding process, additional patterns emerged within the row of each individual student's responses, such as Line 18, which exhibits a clear pattern of simply clicking down the row - with all Equipment post- responses being identical - which at face value would indicate a student who viewed themselves as completely inexperienced at the beginning of the semester and as a total expert at the end of the semester without being aware of any growth in that time.

Line 30 provides an interesting example of a response determined to be unreliable from a student whose short answers indicate they have a solid grasp of the content "A 3d printer uses an STL file and coverts it ingo a Gcode and uses that G code to print", though all Equipment Pre- responses were either 1 or 2, all Reflect responses were 3, and all Now responses were 6; all Competency 4 responses were <Growth without Growth>.

Some of the "potentially unreliable" responses identified in the statistical methods were found to be reliable using the manual coding method. Line 33 provides the clearest example of this; this is the same student response set that is highlighted in Figure 2 as having "two zeros in AAD post columns." Recall that AAD is looking for variance, not patterns; Manual Coding is all about looking for patterns. Given the nature of this group assignment, which was specifically crafted to require students to work with technologies they likely had no prior experience with and where students are required to divide tasks among group members, it is well within the range of patterns one might expect to see that a student who entered the semester with no experience using any of the makerspace equipment might experience tremendous growth in self-perceptions of competence in those areas they used extensively while perceiving no growth in other areas.

Going Forward

As the Likert scales are numerical in nature, it is our hope that an automated script can be developed (or discovered) to facilitate coding tri-set responses in this way. It would be ideal to bring this level of nuance to all Maker Literacies courses' data analysis, though continuing with such an effort using the manual means exhibited here is not an enticing prospect.

Table 2 Manually Coded Competency 4 Tri-Set Responses from IE 4340

	FH	FI	FJ	FK	FL	FM	FN	FO	FP	FQ	FR	FS
1	Pre. C4a	Reflect C4a	Now C4a	Pre. C4b	Reflect C4b	Now C4b	Pre. C4c	Reflect C4c	Now C4c	Pre. C4d	Reflect C4d	Now C4d
2												
3	3	3	3	4	2	3	4	2	3	3	3	3
4	3	3	3	2	2	2	3	2	2	2	2	2
5	1	3	3	1	2	3	3	3	3	3	3	3
6	4	3	4	3	2	3	4	2	3	3	2	3
7	4	3	3	4	4	4	4	4	4	4	4	4
8	1	1	4	1	4	4	3	4	4	3	4	4
9	3	4	4	3	3	3	3	3	3	3	3	3
10												
11	4	3	3	4	3	3	4	4	3	4	3	3
12	4	2	2	3	2	2	3	3	3	3	3	3
13	3	5	5	3	5	5	5	5	5	3	5	5
14	3	3	3	3	2	3	2	4	3	3	3	4
15	3	3	3	3	1	2	3	1	3	3	2	3
16	3	4	4	4	4	4	3	4	4	3	4	4
17	3	4	4	4	4	4	4	4	4	4	4	4
18	3	5	5	1	5	5	1	5	5	3	5	5
19	2	3	3	2	1	1	3	3	3	2	2	2
20	4	3	3	3	3	3	3	4	4	4	4	4
21	3	3	4	2	4	3	3	4	3	3	3	3
22	3	3	4	3	3	4	2	1	3	2	1	3
23	1	1	3	1	1	5	2	1	4	3	1	3
24	4	3	3	3	2	3	3	2	3	3	2	3
25	3	3	3	4	3	4	4	3	4	3	3	4
26	3	3	4	3	4	4	4	4	4	4	3	4
27	3	5	5	3	5	5	3	5	5	3	5	5
28	4	4	4	4	4	4	4	4	4	4	4	4
29	3	2	3	3	2	3	3	1	3	3	1	3
30	3	4	4	3	4	4	3	4	4	3	4	4
31	2	2	3	2	1	3	2	2	2	1	2	2
32	3	3	5	3	3	4	4	3	5	3	3	4
33	1	1	3	1	1	2	1	1	1	1	1	1
34	3	4	3	4	2	4	4	2	4	3	2	3
35	3	1	3	3	1	4	3	1	3	3	1	4
36	4	3	4	4	4	4	5	4	4	4	3	4
37	3	2	3	3	2	3	2	1	3	3	2	4
38	5	3	5	4	3	5	5	3	5	5	3	5
39	3	2	2	2	2	3	2	3	3	2	3	3
40	3	2	4	4	1	4	4	2	4	3	2	3
41	4	3	3	2	3	2	4	3	2	4	4	2
42	3	4	4	3	3	3	5	4	4	5	5	5
43	3	2	3	3	2	3	3	2	3	3	2	3
44	1	3	3	1	3	3	2	4	4	2	4	4
45												
46	3	3	3	3	3	3	3	3	2	4	3	3
47	4	4	5	4	3	4	5	4	4	3	3	4
48	4	4	4	4	4	4	3	4	4	3	4	4
49	5	3	3	4	3	3	5	3	3	4	3	3
50	5	4	5	4	4	5	4	4	4	4	4	4
51	5	5	5	4	5	5	5	5	5	4	5	5
52	5	3	4	4	2	5	5	5	5	4	2	5
53	3	4	3	3	3	3	4	3	3	3	3	3
54	3	3	3	3	2	4	4	4	4	3	3	3
55	4	4	5	3	3	4	4	3	4	3	4	5
56	5	5	5	4	4	4	4	4	5	3	4	4

Table 3 Manually Coded Makerspace Equipment Tri-set Responses; IE 4340

	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP
1	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now	Pre-	Reflect	Now
2	3D printing			3D scanning				Lasers			Vinyl		Screen Printing			Sewing/Serging			CNC Embroidery			CNC Milling		
3	5	5	5	2	2	2	2	3	3	1	1	2	1	1	2	1	1	1	1	1	2	1	1	1
4	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	5	4	4	1	1	1	5	3	4	1	1	1	1	1	1	1	1	1	1	1	1	5	3	3
6	1	2	2	1	2	2	1	2	2	1	2	2	1	2	3	1	4	6	1	2	4	1	2	2
7	5	3	3	5	3	3	4	2	2	4	2	5	4	2	5	4	1	1	4	1	1	4	1	1
8	2	2	6	1	1	1	2	5	5	2	2	2	1	2	2	1	1	6	1	1	6	1	3	3
9	4	5	5	1	1	1	2	1	3	2	1	1	2	1	1	1	1	3	1	2	5	1	1	1
10																								
11	5	6	6	3	4	4	5	6	5	4	5	6	3	5	4	3	4	3	6	7	6	1	5	2
12	2	2	2	2	1	1	2	1	2	2	1	1	2	1	1	2	1	1	2	1	2	1	1	2
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	2	2	1	1	1	1	1	1	1	1	1	1	4	4	1	1	1	1	1	1	1	1	1
15	2	2	2	2	1	1	1	1	1	1	1	3	2	1	5	1	1	1	1	2	1	1	1	1
16	1	1	1	1	1	1	1	1	1	4	1	1	2	3	3	1	1	1	4	1	1	7	7	7
17	1	2	2	1	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
18	1	7	7	1	7	7	1	7	7	1	7	7	1	7	7	1	7	7	1	7	7	1	7	7
19	2	2	4	1	1	1	1	2	3	1	1	1	1	1	1	1	1	4	1	1	3	1	1	1
20	1	2	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4
21	5	5	5	2	4	4	3	3	3	5	6	6	2	3	2	5	4	4	3	6	6	2	2	2
22	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3
23	1	1	5	1	1	2	1	1	6	1	1	5	2	1	3	1	1	2	1	1	3	1	1	2
24	2	2	4	2	1	2	2	1	3	2	1	1	2	1	1	3	1	1	4	2	1	3	4	1
25	7	6	7	4	5	5	3	4	4	1	1	2	1	1	2	3	1	2	1	1	2	1	2	5
26	2	4	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	7	7	4	3	3	4	1	1
28	3	4	4	1	4	4	1	1	3	1	1	1	1	1	1	3	5	5	1	4	5	1	1	3
29	2	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	2	4	1	2	4	1	1	2
30	2	3	6	2	3	6	2	3	6	1	3	6	2	3	6	1	3	6	1	3	6	1	3	6
31	2	2	4	2	3	2	1	1	1	1	1	1	1	1	1	2	4	1	1	1	1	1	1	1
32	2	3	5	1	2	5	1	2	5	1	2	5	1	2	6	2	2	4	1	1	5	1	2	5
33	1	1	5	1	1	1	1	1	1	1	1	4	1	1	4	1	1	2	1	1	1	1	1	1
34	5	3	5	3	1	2	1	1	1	1	2	3	1	1	3	1	1	1	1	1	4	3	1	2
35	4	4	4	1	2	2	1	1	4	1	1	4	1	1	4	1	1	1	1	1	1	1	1	1
36	2	4	5	1	1	1	2	1	2	1	1	2	1	1	2	5	4	5	2	1	2	1	1	2
37	2	3	4	1	1	1	4	5	5	1	1	1	1	1	3	4	2	5	1	1	1	1	1	1
38	2	2	6	1	1	4	1	2	4	1	2	7	2	1	7	2	1	3	1	1	3	1	1	4
39	5	4	4	2	4	4	2	3	3	2	4	4	2	1	1	2	1	1	2	1	1	2	3	3
40	5	5	5	3	1	1	3	1	1	3	1	4	3	1	4	2	1	5	2	1	6	3	2	2
41	2	3	7	2	3	6	2	3	4	2	1	5	2	3	2	3	1	5	2	5	2	2	5	1
42	3	5	5	1	1	1	3	5	5	1	1	1	1	1	1	3	4	4	1	1	4	1	1	5
43	2	1	4	1	1	1	1	1	4	1	1	1	1	1	5	1	1	4	1	1	1	1	1	1
44	2	5	5	2	1	1	2	4	5	1	1	1	1	1	1	1	1	1	1	1	1	1	5	1
45																								
46	2	1	2	1	1	1	1	1	2	1	1	3	1	1	2	2	2	3	1	1	2	1	1	1
47	2	2	4	1	1	1	2	4	3	1	2	2	1	3	3	3	4	5	1	3	4	1	3	3
48	1	4	4	1	4	4	1	4	4	1	4	4	1	4	4	1	4	4	1	4	4	1	4	4
49	1	1	2	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
50	1	2	4	1	3	1	2	1	1	1	2	2	4	2	4	2	1	1	3	1	1	3	1	1
51	3	2	4	2	1	1	2	1	2	3	2	3	2	1	4	5	6	6	2	2	4	2	1	1
52	3	2	3	2	1	1	2	2	1	2	1	1	1	1	2	4	1	5	1	1	4	2	1	1
53	1	2	2	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	2	1	1	1	1
54	3	4	4	1	1	4	2	1	1	1	1	3	1	1	3	2	1	1	1	1	1	2	1	1
55	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1
56	2	4	5	2	2	2	2	1	2	1	1	2	1	1	2	2	3	5	1	1	3	2	1	1