AC 2011-1189: GRAPHICAL COMMUNICATIONS: A CONCEPT INVENTORY

Heidi M Steinhauer, Embry-Riddle Aeronautical Univ., Daytona Beach

Assistant Professor of Freshmen Engineering at Embry Riddle Aeronautical University. Have Introductory Graphics for the last 10 years in addtion to developing several upper level advanced 3D modeling courses. Areas of reserach interest: development of student visualization skills, effective integration of 3D modeling into graphics communication courses, and women's retention in engineering. Co-Advisor of only all-women's baja SAE team in the world and Director of summer engineering camp for middle school girls.

Graphical Communications: A Concept Inventory For Pre/Post Test Assessment

Abstract

Engineering graphics is a fundamental skill required for students to be successful engineers; in the past, a year-long, two-course sequence was devoted to teaching engineering graphics. Today most engineering programs allot a single semester to teaching engineering graphics, in spite of the increased content. Coupled with this decrease in class time is an upturn in student misconceptions, faculty increasingly encounter students who are to correctly solve problems while possessing the incorrect conceptual frame for the posed problems Therefore, engineering graphics faculty must be efficient in their delivery of course material while ensuring that students are learning *and* understanding the key concepts to engineering graphics.

In order to enable faculty to identify key areas of student misconceptions in graphics, we have developed a concept inventory for engineering graphics encompassing line-types, isometric and orthographic projection, scaling, section views, auxiliary views, and dimensioning. A concept inventory is a highly structured standardized test designed to reveal concepts of high importance along with student misconceptions about a specific body of knowledge. In addition to revealing student misunderstandings, concept inventories are now accepted as an assessment tool for ABET as proof of student learning and achievement.

To ensure validity of the concept inventory, a Delphi study was conducted with all the expert faculty members who teach Introductory Graphical Communications, IGC, at Embry Riddle Aeronautical University. The resultant pre/post-test assessment focuses on the conceptual visualization of the theoretical skills of the course. The five experts were asked to individually generate a list of core concepts students need to know to be successful at graphical communications. The initial list of concepts was consolidated and condensed into a preliminary concepts ranking table associated by general concept. The experts were asked to rate all provided concepts in the preliminary ranking table as essential (E), nonessential (N), or optional (O). Experts were also asked to rank each concept as to the difficulty in student learning (0-7). Results from the preliminary concept ranking table, depending on expert congruence and agreement, were used to generate the concept inventory for the pre/post-test assessment. Finally, data from student think aloud activities were used to reword questions and to create the multiple-choice distractors.

This paper focuses on the development and progress of the graphical communications concept inventory focusing on the process and the challenges encountered while creating a new concept inventory, the validity, reliability, item discrimination, item difficulty, and the item total score correlation of the 200 participant field study will be presented. Preliminary findings from the January 2011 field test will also be presented.

Clearly a concept inventory for graphical communication is needed; the Delphi method provided such an instrument that can be used across multiple settings, universities, classrooms. This

concept inventory will help faculty *and* students learn more effectively in the current learning environments provided for engineering graphics.

Context and Purpose Statement

In 2000, ABET, (originally named Accreditation Board for Engineering Technology), issued its revised objectives for the accreditation of undergraduate engineering programs. These revised objectives required a fundamental shift from a numerical list of minimums (i.e. number of hours teaching, number of PhD faculty in department, etc.) to an open list of objectives that each institution could then tailor to their specific programs. Institutions are also now required to provide assessments that "identify, collect, and prepare data to evaluate the achievement of program outcomes and program educational outcomes" (p. 2) [1]. Program education objectives are the overarching statements that define the career and professional achievements expected of the program graduates, and program objectives are the more narrowly defined descriptions of what students are expected to know and be able to do by graduation – the skills, knowledge, and behaviors acquired [1]. Institutions are to formatively evaluate the data and evidence generated by the assessments to continuously improve their programs.

The four ABET 2000 outcomes that directly relate to the freshman graphical communication class are:

a) an ability to apply knowledge of mathematics, science, and engineering
e) an ability to identify, formulate, and solve engineering problems
g) an ability to communicate effectively
k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

A concept inventory that has reported validity and reliability can significantly contribute to the ABET required assessment of program goals. The gathered results may be used to both formatively assess the learning gains of individual students as well as to summatively assess and compare the performance across all sections of the course.

The purpose of this research project is to develop a concept inventory that measures student understanding of the concepts that comprise an introductory graphical communication course. It is expected that students with weaker understanding of these concepts will not score as well on the inventory as those with a stronger understanding. Lower performing students are at a higher risk of becoming discouraged with engineering and subsequently withdrawing from or failing out of the program. Higher performing students are at risk of becoming bored with and disengaging from engineering. This concept inventory could aid educators and administrators in placing students in the appropriate course per their skill level by accurately identifying the lower, the middle, and the higher performing students. The inventory may also be used as a pre and post-assessment for the middle and lower performing students to demonstrate their learning gains. The results of the inventory may also be used to develop new and refine existing lesson plans to address the identified student misconceptions.

Theoretical Framework

Student misconceptions are common throughout the engineering curriculum. Engineering faculty members repeatedly encounter students while able to correctly solve problems possess the incorrect conceptual frame for the posed problem. Faculty observations are supported by evidence in the literature that suggests students do not conceptually understand many fundamental science and engineering theories [2].   The literature also implies the problem is deeper than misinterpretation and confusion; rather, it encompasses a fundamental misconception [12].

To be able to address and correct students' misconceptions, faculty members must first be able to accurately identify them. A concept inventory can aid faculty members in identifying the misunderstandings. Concept inventories are multiple choice tests in which the correct answer *and* the distracters have been thoroughly researched. Concept inventories have been developed for selected sciences: mathematics, statics, force concept for physics, chemistry, thermal and transport science, but few have been developed for engineering sciences – in particular engineering graphics [3, 4].

Engineering graphics teaches a fundamental skill, the language of engineering, required for students to be successful engineers [5-7]. In the past, engineering graphics was taught as a one year, two course sequence. This allowed faculty enough time to identify and correct theoretical and practical misconceptions and afforded students enough time to develop their weak spatial abilities. As the engineering curriculum changed, most colleges and universities have compressed the two course sequence into one course or have even further reduced it to a module embedded within another course. Aggravating this cognitive overload is the addition of 3D computer graphics and modeling. Spatial abilities as applied in graphical communication are not easily learned and many students require significant time and guidance to develop these skills properly.

3-D spatial abilities are not a singular construct, but rather a collection of specific skills – spatial visualization and spatial orientation [8-14]. Spatial ability is the over-arching term encompassing: mental rotation, mental transformation, spatial orientation, and spatial visualization. Figure 1 shows the structure of these interdependent concepts.   Spatial ability involves mentally moving an object.   It is the ability to manipulate an object in an imaginary 3-D space and create a representation of that object from a new viewpoint or perspective.

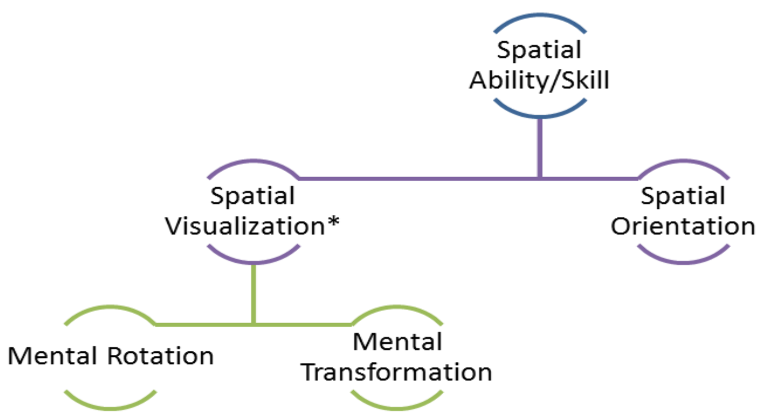

Figure 1 - Spatial Visualization Classification Structure (McGee 1979)

Spatial visualization is comprised of two skills – Mental Rotation and Mental Transformation. Mental rotation is the skill of transforming the entire object by turning it in space. Mental transformation is the skill of altering only a portion of the object. Spatial orientation is the ability to mentally move your viewpoint while the object remains fixed in space. Strong and Smith (2002), offer two examples of spatial orientation: a diver even though she may be turning and twisting, knows exactly where the water is; and a stunt pilot knowing where the ground is during his maneuvers. The concept inventory will be examining the spatial visualization branch of the hierarchy only, as identified by the asterisk.

The ability to visualize objects and situations in one's mind, and to manipulate those images, is a cognitive skill vital to many career fields, especially those requiring work with graphical images. By one estimate, there are at least 84 different careers for which spatial skills play an important role [12]. Spatial ability is an essential skill for Science, Technology, Engineering and Mathematics (STEM) fields. Spatial abilities have been widely studied and are known to be fundamental to higher-level thinking, reasoning and creative processes [3, 4, 12, 13, 15-25]. For technical professions, such as engineering, robotics, and GIS (Geographic Information Systems), operate laparoscopic equipment [26], and to interact efficiently with database management [27], spatial visualization skills and mental rotation abilities are especially important [18, 28]. Ferguson [29] believes that several well-known engineering failures: the Challenger explosion, the Columbia disaster, the Hubble space telescope, and the USS Vincennes Aegis system, have occurred due to the removal of visual, tactic, and sensory applications from the current engineering curriculum.

Students struggle with the transition from 2D to 3D and back to 2D, only with time and repeated structured guidance are students able to properly develop these skills [4]. Given the limited course time now available, engineering graphics faculty must be efficient in their delivery of the course material to ensure that students are learning *and* understanding the key concepts to engineering graphics. A standardized engineering graphics concept inventory would allow for faculty to identify key areas of student misconceptions.

Traditionally students' spatial abilities were improved as a by-product of the engineering graphics curriculum and not due to any formally addressed outcomes. Ferguson [29] claims that the engineering education of today has diverged too much from its artistic, visual beginnings, and that the curriculum relies too heavily on analytical methods and not enough on tactile and visual perception. With the integration of the ABET 2000 criterion a-k there has been a shift in engineering education to outcomes based assessment [30-32]. One of the frequently cited educational outcomes of engineering design graphics is the development of student's 3-D spatial visualization abilities. During the last decade there have been some changes in the contents of the Engineering Graphics discipline, but barring some exceptions, spatial abilities are still considered as a secondary goal that simply is achieved through the learning of other concepts [9].

Since the release of ABET EC 2000 – specifically criterion 3a-k, engineering educators have searched for appropriate methods to provide supporting data that they are meeting the requirements for student learning and outcomes. A validated concept inventory, implemented as a pre and post-test assessment, would provide much of the data required to show an engineering graphics course is meeting established ABET student outcomes of a, g, and k.

The instrument is to be used to measure and establish a baseline of student misperceptions of the application of visualization skills in engineering graphics. These abilities are a fundamental skill required of all engineering students. Students who have these skills are able to decompose a 3-dimensional (isometric) object into the proper 2-dimensional (orthographic) views while properly dimensioning those orthographic views. Students who have very weak to no visualization skills are not able to properly complete these tasks.

Existing Assessment Instruments

There are several well-known assessments of spatial visualization, the Purdue Spatial Visualization Test, the Mental Rotation Test, and the Mental Cutting Test. While these three are by no means the only ones, they are well known and have been widely utilized. A brief description of these three tests is provided.

Purdue Spatial Visualization Test: Rotations

The PSVT:R was developed by Guay [33] at Purdue University. In this timed 30 item test participants are asked to visualize the direction and rotation of the provided sample model , next they are to visualize what the second object would appear as if it were rotated the same as the sample, and then they are to select the correct image from the five options provided. A sample question from the PSVT:R is provided in Figure 2.

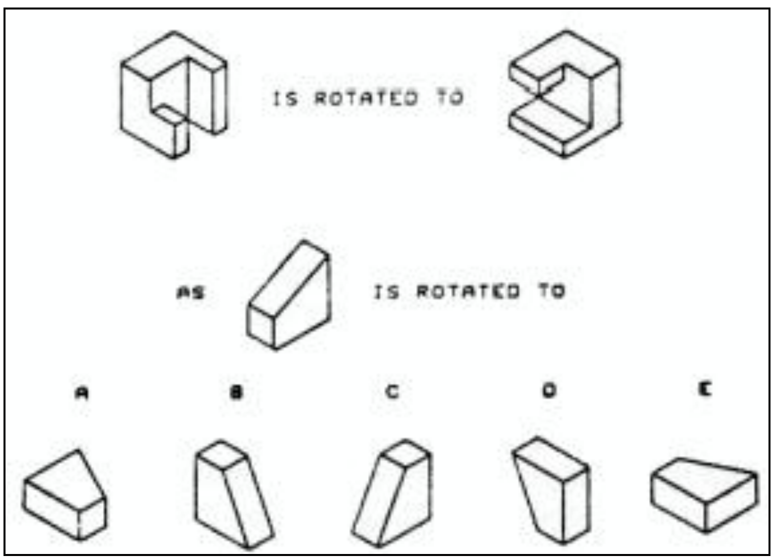

Figure 2 – Sample PSVT:R Question

Mental Rotation Test, MRT

The MRT, Mental Rotations Test, was developed by Vandenberg and Kuse [34]. The figures they used were derived from Shepard and Metlzer's earlier work [35]. The MRT is a timed 20 item test divided into sets of four test items. Each item contains the criterion model, two correct representations of the criterion object, and two distractors. Participants are asked to identify the two correct representations. A sample question from the MRT is provided in Figure 3.

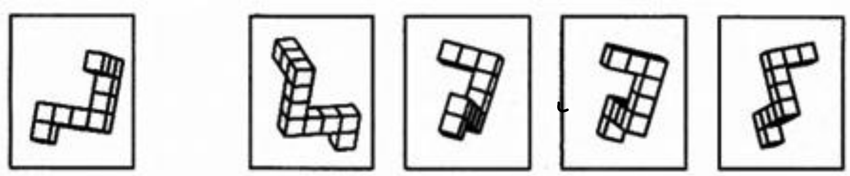

Figure 3 – Sample MRT Question

Mental Cutting Test, MCT

Developed in 1939 as a college examination, the CEEB Special Aptitude Test in Spatial Relations has been used for many years to determine spatial ability. The MCT is a timed 25 problem test divided into two categories: pattern recognition and dimension specification. For the former type of problem only the pattern of the section needs to be identified to determine the correct answer, for the latter both the pattern and the correct dimensions need to be identified.

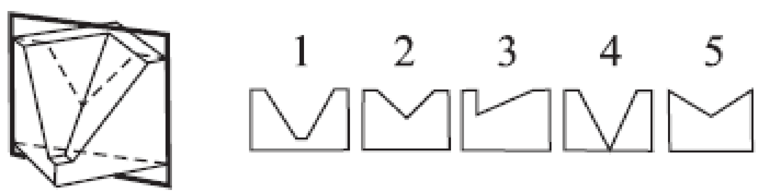

Figure 4 - Sample MCT Question

The reported reliability of the PSVT:R, KR-20 was .87, .89, and .92 for university students [36-38] and the reported reliability of the MRT, Chronbach's α, of the MRT was .87 and .92 for undergraduate chemistry and engineering students [39, 40]. While these tests have a long history of reported reliability they are not measuring the direct concepts of engineering graphics – rather they are focused on instead spatial visualization. A concept inventory on engineering graphics would better identify the specific areas of weakness.

Aspects of Validity

Messick [41] brought into question the traditional view of validity where it is neatly divided into three distinct categories: content, criterion, and construct, instead he posited there is one type validity – construct validity which is supported by six types of construct validity evidence. He argued that traditional approach does not consider the "value implications of score meaning as a basis for action and the social consequences of score use," [41] page 741. Instead he posits that validity is a unified concept that interrelates the three into a framework to guide empirical testing. Messick [41] lists the six evidences of construct validity as: content, substantive, structural, generalizability, external, and consequential. Content validity is defined by the range of knowledge, skills, abilities, etc. to be examined by the test instrument. Substantive validity is described as an appropriate sampling of the tasks required to demonstrate suitable coverage of the domain. Structural validity refers to the organization of the scoring and that it is inherently related to the domain under investigation. The generalizability aspect is to ensure that the results are applicable to the entire domain and not just to the area tested. The results are said to have external validity if they can be directly compared or correlated to another instrument from a similar domain. Finally, the consequential facet of validity entails the consequences of using the scores as a source or call for action and the probable impact of doing so.

Validity Evidence

Throughout this research project Messick's framework was used to evaluate the validity of the developed instrument. Each of the six aspects of validity have been carefully evaluated. Each of the six aspects and how they have been addressed will be covered in detail in this section.

Content validity was addressed through two methods: 1) the use of a clearly defined purpose statement and 2) the use of experts to generate and review test items. For this study, experts were defined as educators who have taught this course for at least five years.

Substantive validity was accounted for through multiple approaches: 1) the developed construct map is supported by current research, 2) the operational definition is consistent with the construct map, 3) the internal model follows the construct map and operational definition, 4) the external model is supported by the current research, and 5) the items and distractors were refined and developed by think-alouds methods and debriefing interviews with former students.

Structural validity has been addressed through the alignment of the internal model and the structure of the instrument.

Generalizability has been supported by the selection of the target population – college level freshmen engineering students. The items are expected to have the same characteristics across the varied subsets of the target population.

External validity will be addressed through the multiple field tests of the instrument. The instrument was field tested in December 2010, and will be field tested again in January 2011. It is expected these results will demonstrate that the individuals have changed (improved) over time.

Finally, consequential validity refers to how the scores of the instrument will impact the participants who do not pass or perform well. While this is not currently an issue, it will be when the instrument is finalized. The test is to be criterion based with a cut score of 10 out of a possible 16, this cut score was determined by the panel of experts who helped to the develop the test items. It is important to note that not passing the inventory merely points to the specific areas of spatial visualization to be address during the course of the semester. Students are not to be informed of their individual earned scores; rather they will be reported in the aggregate form for the entire class.

Preliminary Findings From the Pilot Studies

Operationalization of Spatial Abilities

Spatial abilities are comprised of two primary factors: spatial visualization and spatial orientation. Spatial visualization is comprised of two skills: mental rotation and mental transformation. The items of potential indicators for the spatial ability inventory include: selfassessment of skills and instructor feedback and observations. The content sampling for this inventory will be developed using instructor feedback and observations. The four categories of spatial skills to be investigated will be: surface prediction, section views, auxiliary views, and dimensioning.

The inventory will have both multiple choice and Likert items. The Likert items will be a three point scale, spanning from "highly confident" to "not confident at all." The ordinal nature of the Likert scale allows for the responses to be numerically coded, with "highly confident" assigned three points and "not confident at all" assigned zero points. These responses can be averaged to provide a score that can be mapped back to a respondent's skill confidence. The use of these scales along with the think aloud protocols during the distractor development, and the feedback provided by the expert panel all serve to strengthen the substantive validity. Instead of being norm-referenced instrument, which would only indicate whether the individual's performance was better or worse, this instrument will be criterion-referenced. The test score can be used to interpret the relationship between the individual's knowledge and the appropriate or required skill level. Structural validity will be addressed through reliability testing, the ability for similar to "hang" together, based upon those results, inventory items maybe revised, removed, or replaced.

Pilot Testing and Item Analysis – August 2010

Following the item analysis, the reliability, α, of the inventory is .47; the generally accepted minimum for reliability is around .70. Several of the inventory items had a low reported reliability, it is apparent these items of the inventory will be need to revised, before this instrument can be formally implemented as an ABET program objectives assessment.

There were a total of 217 students who participated in the pilot study. Although 217 participants completed the inventory, due to missing data responses from only 172 were included in the item analysis. Participants responded to 16 multiple choice questions. The concept inventory was administered as a paper based test, there were no time restrictions placed on the participants. JMetrik was used to conduct the item analysis. The instrument had a reliability of .47, due to the lower reliability of the results, several items will need to be reworded and revised. These items are discussed in detail in following section. The test level descriptive statistics are included in Table 1. The results of the item analysis and recommendations for each test item are discussed in the following section.

Table 1 – August 2010 Test Level Descriptive Statistics

# of Items	# of Participants	Mean	Median	Standard Deviation	Minimum	Maximum
16	172	6.4070	6.000	2.4892	1	12

Of the 16 items in the concept inventory only the items requiring revision are discussed in following section, all revisions are explained in detail, with the original distractors along with item totals, and the revised distractors included for each of these items.

Item 2 – If Surface A was an inclined surface it would appear as what in the Right Side View

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.3488	.4780	.0292	.4715	.4866	Revise

Based on the item difficulty and the bi-serial correlation with total, the original question for item 2 has been revised for the final inventory. Upon reviewing the item totals of the four distractors, it was decided the distractors needed to be significantly reworded. The original five distractors along with their item totals are listed below.

A (.4360), an inclined line
B (.1047), a surface
C (.0581), a horizontal line
D (.3488)*, either A and B
E (.0465), none of the above

The question's distractors were poorly written and confusing; this was indicated by several factors. Both A and B were correct answers and option D included both of these, however the item totals demonstrated the students did not recognize this, as 54% choose either distractor A or B, only 35% selected the correct answer D, and 5% selected C and E respectively. The original and revised distractors and correct answers are included below:

Original	Revised
A: an inclined line	A: a surface or a vertical line
B: a surface	B: a surface or a horizontal line
C: a horizontal line	C: a surface or an inclined line*
D: Either A or B*	D: an inclined line or a vertical line
E: None of the Above	

The correct answers are indicated by an asterisk.

Item 4 – If Surface C was an oblique surface it would appear as what in the Top and Right Side Views?

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.2093	.4080	-.0665	.4715	.4957	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 4 has not been deleted, but distractor C has been reworded for the final version of the inventory. 39% of the students selected distractor C, 23% selected distractor B, 17% selected distractor A and only 23% selected the correct answer. Distractor C was originally worded as "an angled line in one and a surface in the other." It was determined that the inclusion of inclined edge was confusing to students as oblique surfaces are a specific type of inclined surface, to eliminate this confusion inclined edge was replaced with vertical edge.

Item 5a – Based on the part's geometry the most appropriate section view would be:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.4186	.4948	.0174	.4715	.4898	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 5a has not been revised and kept in its original form for the final inventory.

Item 5b – Because

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.2267	.4199	.0394	.4715	.4818	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 5b has not been deleted, but distractor E has been removed and distractor F has been reworded for the final version of the inventory. 11% of the students selected distractor E, 39% selected distractor F, and only 23% selected the correct answer. Upon close examination of the provided part (which is not symmetrical) and a discussion with other experts it was determined that students are quite familiar with the concepts of symmetry and non-symmetry, so it was decided to eliminate those distractors. Original distractors along with the item totals are included below.

E: (.1105), the part is symmetrical
F: (.3837), the part is not symmetrical
G: (.2267)*, the part contains non-planer internal details
H: (.2733), the part contains an inclined surface

The correct answer is identified by an asterisk, the distractor to be removed is E, and distractor F has been reworded. The revised distractors are listed below.

E: the part is contains internal circular details
F: the part contains non-planer internal details*
G: the part contains an inclined surface

The correct answer is identified by an asterisk.

Item 6 – The most appropriate view for generating an auxiliary view for Surface A is:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.1395	.3475	.1139	.4715	.4708	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 6 has not been deleted, however the distractors have been revised for the final version of the inventory. Question 6 originally had one correct answer (B) and five distractors. Due to a formatting/editing error the last two distractors (E and F) were mistakenly included; these have been deleted from the final version of the concept inventory.

Item 7 – When dimensioning an engineering drawing, primary units are in inches, select the appropriate series of dimensions:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.4302	.4966	.2518	.4715	.4443	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 7 will not be deleted, however distractor E (none of the above) has been removed for the final inventory, as only 9% selected it as a distractor. The original distractors and their Item totals are included below.

A (.1105), 1.5; 2.0; 3.8; and 4.3
B (.4302)*, 1.50; 2.00; 3.75; and 4.25
C (.2209), 1 ½; 2; 3 ¾; and 4 ½
D (.1395), 1.5; 2; 3.75; and 4.5
E (.0988), none of the above

The correct answer is indicated by an asterisk and the distractor selected to be deleted is E.

Item 8 – When dimensioning an engineering drawing, primary units are in mm, select the appropriately written half scale:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.5233	.5009	-.0685	.4715	.5064	Revise

Based on the item difficulty, the bi-serial correlation with total, and the Cronbach's α if item is deleted, item 8 will not be deleted, however distractors A and B have been revised for the final version of the inventory, as only 10% of the students selected them. The original distractors and their Item totals and the revised distractors are included below.

Original	Revised
A (.0349), 2 =1	A: 2/1
B (.0756), 1 = 2	B: 1/2
C (.3663), 2 :1	C: 2 : 1
D (.5233)*, 1: 2	D: 1 : 2*

The correct answers are indicated by an asterisk.

Item 9a – Dimension A in the Front View is incorrect because:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.1686	.3755	-.1017	.4715	.4967	Revise

Based on the item difficulty and the bi-serial correlation with total, the original question for item 9a has been revised for the final inventory. Upon reviewing the item totals of the four distractors, it was decided the distractors needed to be significantly reworded. The original five distractors along with their item totals are listed below.

A: (.1105), the dimension is dimensioning an element that is not true size
B: (.2500), the dimension is a duplicate dimension and not required
C: (.1686)*, the dimension is not placed in the most descriptive location
D: (.1628), the dimension format is incorrect for dimensions that are in inches
E: (.3028), the dimension as provided is correct

The question's distractors were poorly written and confusing; this was indicated by several factors. C was the correct answer, however the item totals demonstrated the students did not recognize this, as 30% chose distractor E, 25% choose either distractor B, only 17% selected the correct answer D, and 11% selected A. Distractor B could have been interpreted as a correct option if students did not properly understand the provided image, based on this distractor B has been eliminated. The revised distractors and correct answer is included below:

Revised

A: the dimension is dimensioning an element that is not true size
B: the dimension is not placed in the most descriptive location*
C: the dimension format is incorrect for dimensions that are in inches
D: the dimension as provided is correct

The correct answers are indicated by an asterisk.

Item 10a – Dimension B in the Right Side View is incorrect because:

Item Difficulty	Standard Deviation	Biserial Correlation with Total	Cronbach's α	Cronbach's α if Item is Deleted	Decision About Item
.1806	.3903	.0240	.4715	.4822	Revise

Based on the item difficulty and the bi-serial correlation with total, the original question for item10a has been revised for the final inventory. Upon reviewing the item totals of the six distractors, it was decided several distractors needed to be significantly reworded and additional one needed to be deleted. The original six distractors along with their item totals are listed below.

A (.0581), the dimension is dimensioning an element that is not true size
B (.1628), the dimension includes information that is not needed
C (.2849), the dimension is not placed in the most descriptive location
D (.0814), the dimension as provided is correct

E (.0814), both A and B
F (.1860)*, both A and C
G (.1458), both B and C

The question's distractors were poorly written and confusing; this was indicated by several factors. Both A and C were correct answers and option F included both of these, however the item totals demonstrated the students did not recognize this, as 33% choose either distractor A or C, only 18% selected the correct answer. The revised distractors and correct answer are included below:

A: the dimension is dimensioning an element that is not true size*
B: the dimension includes information that is not needed
C: incorrectly combines three dimensions into one
D: none of the above, the dimension as provided is correct.

The correct answers are indicated by an asterisk.

The item analysis resulted in a reliability coefficient ($\alpha = .47$), item difficulties ranged from .1395 to .6047, and the biserial correlation with totals ranged between -.1017 and .6120. The reported numbers are not ideal and indicate the concept inventory needs to be revised and retested before formal implementation of a program's goals, there are some good indicators. This concept inventory is to be criterion based and therefore the biserial correlation need only be positive, only three of the 16 items had a negative value. These three items have been revised to improve the negative values. The item difficulty were well distributed between less challenging to more challenging with more items placed towards the less challenging end of the spectrum: 3 items were in the .10-.20 range; 2 items were in the .21-.30 range; 6 items were in the .41-.50 range; and 5 items were in the .51-.61 range.

Pilot Testing and Item Analysis – January 2011

The reliability, α, of the inventory is .67; the generally accepted minimum for reliability is around .70. While still a bit low it is a marked improvement from reliability of the August 2010 pilot study. This improvement to reliability is supported by the fact that only two of inventory items are marginally negatively impact the overall instrument reliability.

There were a total of 125 students who participated in this pilot study. Only one inventory needed to be removed due to missing data. Participants responded to 16 multiple choice questions. The concept inventory was administered as a paper based test, there were no time restrictions placed on the participants. JMetrik was used to conduct the item analysis. The instrument had a reliability of .67. Due to the marginal impact of the two negatively scoring items the inventory will not be revised. The inventory will be field tested a final time in August 2011. This final field study will be implemented at several institutions.

The test level descriptive statistics of the January 2011 field test are included in Table 2.

Table 2 - January 2011 Test Level Descriptive Statistics

# of Items	# of Participants	Mean	Median	Standard Deviation	Minimum	Maximum
16	124	7.3226	7.000	2.7246	0	16

The item analysis resulted in a reliability coefficient ($\alpha = .67$), item difficulties ranged from .1290 to .6371, and the biserial correlation with totals ranged between .0785 and .4755. The reported numbers are much better. All 16 items now have a reported positive biserial correlation; the items flagged in the August 2010 field test have been corrected.

The item difficulty were well distributed between less challenging to more challenging with more items placed towards the less challenging end of the spectrum: 2 items were in the .10-.20 range; 1 item was in the .21-.30 range; 2 items were in the .31-.40; 2 items were in the .41-.50 range; and 8 items were in the .51-.60 range; and 1 item was in the .61-.70 range.

Limitations and Future Research

While the low reliability of the August 2010 field study has been addressed with the January 2011 field test, there is still room for improvement. This improvement can be attributed to several factors. First, all four professors were given detailed instructions to follow. The instructions specified the specific date and time to implement the inventory. Professors were asked to give the inventory at the start of the class and to allow students no more than 15 minutes to complete it. The instructions also provided an explanation to the students that the field test was not testing their specific knowledge rather it was testing the instrument itself. Of the 125 concept inventories given 124 were completed, this was vast improvement.

One weaknesses still present is implementing the inventory given during the first class when students have yet to become situated and comfortable with the class structure and environment. Nor were the classroom settings and locations uniform across the sections. A third weakness is the limited implementation of the field study, only one institution was utilized. Finally, the sample size was smaller than hoped; only 124 were analyzed.

A final pilot study will be implemented in August 2011. It is expected that 300 participants will complete the survey at Embry Riddle Aeronautical University. The concept inventory will also be field tested with approximately 1000 participants at Virginia Tech. Finally, the concept inventory will be field tested at Daytona State College, 100 participants are anticipated. This final field study will address two of the major weaknesses – a limited number of participants and the single institution implementation. This final field test will be implemented as a pre/post-test to gage student learning gains; approximately 1400 participants will participate in the field study. The results will be used to determine the validity of the instrument. It is expected the concept inventory will be ready for formal implementation in January 2012 following the data analysis of the final field study.

1. ABET, *Criteria For Accrediting Engineering Programs: Effective Evaluations During the 2010-2100 Accreditation Cycle*, A.B.o. Directors, Editor. 2009.
2. Streveler, R., et al., *Using a Delphi study to identify the most difficult concepts for students to maser in thermal and transport science*, in *American Sociery for Engineering Education Annual Conference*. 2003: Nashville, TN.
3. Demel, J.T., Frank M. Croft, Frederick D. Meyers,, *Development of Nationally Normed Engineering Graphics Concepts and Skills Tests*, in *2002 ASEE Annual Conference and Exposition*. 2002: Montreal, Quebec, Canada.
4. Demel, J.T., Frederick Meyers, and K.A. Harper, *Developing a Nationally Normed Test for Engineering Graphics - First Pilot Tests and Results*, in *2004 ASEE Annual Conference and Exposition*. 2004: Salt Lake CIty, UT.
5. Orth, H.D., *Establishing and maintaining standards of excellence in drawing.* Journal of Engineering Drawing, 1941. **5**(1): p. 7-10.
6. Rising, J.S., *Integration of engineering drawing and descriptive geometry.* Journal of Engineering Drawing, 1948. **12**(3): p. 16-17; 30.
7. Svensen, C.L., *Drawing in engineering education.* Journal of Engineering Drawing, 1948. **12**(3): p. 6-7; 20-21.
8. Olkun, S., *Making connections: Improving spatial abilities with engineering drawing activities.* International Journal of Mathematics Teaching and Learning, 2003. **April**.
9. Contreras, M.J., et al., *Sex differences in dynamic spatial ability: The unsolved question of performance factors.* Memory and Cognition, 2007. **35**(2): p. 7.
10. Voyer, D., M.A. Rodgers, and P.A. McCormick, *Timing conditions and the magnitude of gender differences on the Mental Rotations Test.* Memory and Cognition, 2004. **32**(1): p. 11.
11. Quaiser-Pohl, C. and W. Lehmann, *Girls' spatial abilities: Charting the contributions of experiences and attitudes in different academic groups.* British Journal of Educational Pyschology, 2002(72): p. 15.
12. Smith, I.M., *Spatial ability - Its educational and social significance*. 1964, London: University of London.
13. Sorby, S.A., *Spatial Abilities And Their Relationship To Computer Adided Design Instruction*, in *ASEE Annual Conference and Exposition*. 1999: Charlotte, NC.
14. Strong, S. and R. Smith, *Spatial visualization: Fundamentals and trends in engineering graphics.* Journal of Industrial Technology, 2002. **18**(1): p. 6.
15. Bishop, J.E., *Developing students spatial ability.* Science Teacher, 1978. **45**: p. 3.
16. Branoff, T., *Large course redesign: Revising an Introductory engineering graphics course to move from face-to-face to hybrid instruction*, in *ASEE Annual Conference and Exposition*. 2009: Austin, TX.
17. Gorska, R. and S.A. Sorby, *Testing instruments for the assessment of 3-D spatial skills*, in *ASEE Annual Conference and Exposition*. 2008: Pittsburgh, PA.
18. Hamlin, A.H., N. Boersma, & S. Sorby, *Do Spatial Abilities Impact The Learning Of 3-D Solid Modeling Software?*, in *ASEE Annual Conference*. 2006: Chicago, IL.
19. Hamlin, A.H., N Veurink, & S. Sorby, *Impact Of Spatial Visualization Topic Order On Student Performance And Attitudes*, in *ASEE Annual Conference*. 2008: Pittsburgh, PA.
20. Harris, L.A., & F. Meyers, *Graphics: Into The 21st Century*, in *ASEE Annual Conference and Exposition*. 2007: Honolulu, HI.
21. Sorby, S.A., & M. F. Young, *Assessment of a Visualization-Based Placement Exam for a Freshman Graphics Course*, in *ASEE Annual Conference and Exposition*. 1998: Seattle, WA.
22. Sorby, S.A., *A "new and improved" course for developing spatial visualization skills*, in *ASEE Annual Conference and Exposition*. 2001: Albuquerque, NM.
23. Sorby, S.A., *Integrating graphics in a first-year engineering program*, in *ASEE Annual Conference and Exposition*. 2003: Nashville, TN.

24. Sorby, S.A., *Developing 3D spatial skills for engineering students.* Australasian Assocation of Engineering Education, 2007. **13**(1): p. 11.
25. Sorby, S.A., *Educational research in developing 3-D spatial skills for engineering students.* International Journal of Science Education, 2009. **31**(3): p. 22.
26. Eyal, R. and F. Tendick, *Medicine meets virtual reality 2001*. Spatial ability and learning the use of of an angled laparscope in a virtual environment, ed. e.a. J. D. Westwood. 2001, Amsterdam: IOS Press. 7.
27. Norman, K.L., *Spatial visualization - A gateway to computer-based technology.* Journal of Special Education Technology, 1994. **XII**(3): p. 12.
28. Yue, J., *Spatial visualization by isometric drawing*, in *IJME - INTERTECH Conference*. 2006: New York City, New York.
29. Ferguson, E.S., *Engineering and the mind's eye*. 1992, Cambridge: MIT Press.
30. Barr, R.E., Thomas J. Krueger, & Theodore A. Aanstoos, *Addressing Program Outcomes in a Freshmen Introduction to Engineering Course,* in *ASEE Annual Conference and Exposition*. 2005: Portland, OR.
31. Barr, R.E., Thomas J. Krueger, & Theodore A. Aanstoos, *Assessing Student Outcomes in an Engineering Design and Grpahics Course*, in *ASEE Annual Conference and Exposition*. 2003: Nashville, TN.
32. Barr, R.E., Thomas J. Krueger, & Theodore A. Aanstoos, *Results of an EDG Student Outcomes Survey*, in *ASEE Annual Conference and Exposition*. 2004: Salt Lake City, UT.
33. Guay, R., *Purdue Spatial Visualization Test - Visualization of Rotations*. 1977, Purdue Research Foundation: W. Lafayette.
34. Vandenberg, S.G. and A.R. Kuse, *Mental rotations, a group test of three dimensional spatial visualization.* Perceptual and Motor Skills, Al, 1978: p. 599-604.
35. Shepard, R.N. and J. Metzler, *Mental rotations of three-dimensional objects.* Science, 1971. **171**(3972): p. 3.
36. Guay, R.B., *Spatial ability measurement: A critique and an alternative*, in *American Educational Research Association*. 1980: Boston, MA.
37. Sorby, S.A. and B.J. Baartmans, *A course for the development of 3-D spatial visualization skills.* Engineering Design Graphics Journal, 1996. **60**(1): p. 8.
38. Bodner, G.M. and R. Guay, *The Purdue visualization of rotations test.* The Chemical Educator, 1997. **2**(4): p. 1-17.
39. Caissie, A.F., F. Vigneau, and D.A. Bors, *What does the Mental Rotation Test measure? An analysis of difficulty and item characteristics.* The Open Psychology Journal, 2009. **2**: p. 94-102.
40. Branoff, T.J., *Spatial visualization measurement: A modification of the Purdue Spatial Visualization Test - Visualization of rotations.* Engineering Design Graphics Journal, 2000. **64**(2): p. 14-22.
41. Messick, S., *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning.* American Psychologist 1995. **50**: p. 741-749.