UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

A CONTRIBUTION TO CONVEYING QUALITY CRITERIA IN MECHANICAL CAD MODELS AND ASSEMBLIES THROUGH RUBRICS AND COMPREHENSIVE DESIGN INTENT QUALIFICATION

A Dissertation

by

JEFFREY MARTIN OTEY

Submitted to the Doctoral School of

Universitat Politècnica de València

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

under the supervision of

Dr. Manuel Contero González (Universitat Politècnica de València)

Dr. Pedro Company Calleja (Universitat Jaume I)

Major Subject: Design, Manufacturing and Management of Industrial Projects

October 2017

ACKNOWLEDGEMENTS

I would like to express my extreme indebtedness to Dr. Pedro Company Calleja and Dr. Manuel Contero González. Thank you for encouraging me to continue my studies and to take a medical leave of absence after my diagnosis. I appreciate your mentoring, concern, great patience, and guidance more than you know. I am forever grateful. I am also thankful to all my co-authors who contributed to my research and thesis.

I am grateful to Richard Hutchinson and Peter England, who stood by me and assisted me when I returned to work. I am most thankful to my great friend, Jorge Dorribo Camba, who encouraged me throughout this process and was a sounding board for all my ideas and approaches. I could not have done it without you.

Most importantly, I would like to thank my wife and daughter, who stood by me during this process, especially during my treatment and recovery, and gave me unwavering support. I have been blessed beyond all measure.

ABSTRACT

This thesis spotlights the approach of using rubrics to transition basic CAD users to experts, by following an instructional approach designed to produce high quality CAD models. The approach forces a strategic-aimed focus, since the strategic knowledge (as opposed to procedural knowledge) is related to identifying alternate working procedures and how best to choose between them to ensure the most efficient use of the CAD tool.

Appraisal of student performance is a critical component necessary for engaged student learning. The use of rubrics to perform assessment not only serves as a method for instructors to objectively judge student work, but also provides for learner self-assessment.

This research examined the use of assembly rubrics, described how they evolved from parts rubrics, and studied how they affect student self-evaluation. Instructor assessment of students was also evaluated, finding that while the assembly rubrics were partially understood and effectively used by the students, they were more successfully utilized by the instructors.

An exhaustive survey was conducted on how CAD model quality has been historically defined, most often described by examining design intent and conccurrently emphasizing its historical relationship with design rationale. Design rationale has been well-established in industry (in the context of MCAD systems), while design intent remains an often misunderstood concept. Design intent is often defined as a model's anticipated behavior when undergoing alteration, and modeling approaches influence the communication of design intent, with significant advantages to expressing design intent through suitable modeling strategies, especially for CAD beginners.

Strategies designed to improve design intent communication in CAD models, in order to enhance their quality, with guidelines targeted to evaluate efficiency, have been addressed with this research. It is apparent that metrics directed toward the instruction of design intent are needed, since design intent transferred through CAD models can be performed at three stages (sketches, reference datums, and modeling operations), with competing tradeoffs that must be balanced to arrive at the best modeling strategy.

Research included the development of a validation approach that reflects that rubrics are valuable devices to expedite consistent design intent communication, and are vital not only for evaluation, but also for the communication of instructor expectations. This research examined how to clearly define qualities of design intent (among other rubric dimensions) to enable easier CAD assembly assessment.

Consistency was preserved between the parts and assembly rubrics, with variation existing between the descriptors and achievement levels for each criterion, which were restated to be more applicable to assemblies. The use of ambiguous and explicit descriptors were required to prevent indiscriminate evaluation and to ensure proper evaluation. Achievement levels reflected the amount of conformity for each assessed aspect, with simple cases being dichotomously determined. Rubric dimensions were divided into six categories: valid, complete, consistent, concise, clear, and design intent. Annota e-Rubrics platform was used during the experiments.

It has been found that there is more inter-rater agreement and correlation between instructors than between instructors and students, for all rubric dimensions. There is strong to moderate correlation between instructors for the dimensions of validity, completeness, conciseness, and clarity, while slight correlation exists for the dimensions of consistency and design intent. The minor differences between instructors allows the conclusion that the proposed assemblies rubric provides an objective accumulative assessment of students, so it can be assumed that raters can be safely interchanged.

Previous original contributions of the research team were also considered and further improved for this thesis. Firstly, the rubric generated scores, with the scoring process defined by formulas, providing an aggregated score from the achievement levels. The rubric was improved by using three

novel characteristics: criteria dichotomization, evaluation criteria with varying importance levels, and Go/No Go criteria.

Secondly, rubrics can also be described as being either static or dynamic. Static rubrics, existing in paper form only, do not provide immediate feedback to the learner. Dynamic rubrics perform calculations that provide immediate evaluative observations to the user. Besides, they can be independently adapted to specific situations depending on the capability of the user. Electronic rubrics (eRubrics) are ideally suited for dynamic rubrics, and permit the use and development of both adaptable and adaptive rubrics, as described next.

Thirdly, rubrics need to be adaptable (modified by the user), which should make them easily understood and user-friendly, and adaptive (rubric can change itself, depending on the usage pattern). Evaluative rubrics are used when an expert determines the pedagogical progress of a learner, while formative rubrics are employed by the learners themselves, in order to chart their progress and identify scholastic deficiencies that are in need of remediation. Rubrics must be continually refined and improved, in an iterative, collaborative process, until satisfactory agreement is attained, both between raters, but also between raters and learners. Thus, assertions maps were developed, illustrating how the expand-contract strategy adapts the rubrics to CAD trainee progress, while assisting the understanding of the different rubric dimensions. Detailed rubric dimension curricular material (consisting of checklists, good practices, and various evaluation tools) were developed to investigate if they are advantageous for student comprehension, especially when trade-offs to achieve valid designs are needed.

Based on the assembly rubric experiments, it is apparent that the small differences between instructors suggests that the proposed assemblies rubric is sufficiently sophisticated to furnish an unbiased accumulative assessment of student performance. Accordingly, it can be confidently stated that raters can be used interchangeably without sacrificing accuracy. However, the assembly rubric possesses finite efficacy to produce formative self-evaluation of CAD assembly skills for new learners.

Of all of the rubric dimensions, the meaning of design intent has proven to be the most difficult to not only define, but to convey to the students. There are many reasons for this to be so, but primarily, design intent recurrently requires precise prior knowledge of how the mechanism will and should perform, an awareness that may be beyond the comprehension level of inexperienced students.

RESUMEN

Esta tesis pone el foco en el planteamiento de usar rúbricas para transitar de usuarios CAD básicos a expertos, siguiendo un enfoque pedagógico diseñado para producir modelos CAD de alta calidad. El método fuerza un enfoque estratégico, dado que el conocimiento estratégico (al contrario que el conocimiento procedimental) está relacionado con identificar procedimientos de trabajo alternativos y cómo hacer la mejor elección entre ellos para garantizar el uso más eficiente de la herramienta CAD.

La evaluación del rendimiento del estudiante es un componente crítico necesario para el aprendizaje de los estudiantes implicados. El uso de rúbricas para realizar la evaluación sirve no sólo como un método para que los instructores juzguen objetivamente el trabajo de los estudiantes, sino también facilita la autoevaluación del propio educando.

Esta investigación examinó el uso de rúbricas de ensamblaje, describiendo su evolución a partir de rúbricas de piezas, y estudió cómo afectan a la autoevaluación de los estudiantes. También se valoró la evaluación de los estudiantes por los instructores, encontrando que, mientras que las rúbricas de ensamblaje fueron parcialmente comprendidas y utilizadas de manera eficiente por los estudiantes, éstas fueron usadas con más éxito por los instructores.

Se realizó una revisión exhaustiva sobre cómo ha sido definida históricamente la calidad del modelo CAD, muy a menudo descrita mediante el examen de la intención de diseño y enfatizando concurrentemente su relación histórica con la motivación de diseño. La motivación de diseño ha sido bien establecida en entornos industriales (en el contexto de los sistemas MCAD), mientras que la intención de diseño sigue siendo un concepto frecuentemente mal comprendido. La intención de diseño se define a menudo como el comportamiento anticipado del modelo cuando se somete a alteraciones, y los métodos de modelado influyen en la comunicación de la intención de diseño, con ventajas significativas para expresar la intención del diseño a través de estrategias de modelado adecuadas, especialmente para principiantes CAD.

En esta investigación se han abordado estrategias diseñadas para mejorar la comunicación de la intención de diseño en modelos CAD, acrecentando así su calidad, con directrices dirigidas a evaluar su eficiencia. Es evidente que se necesitan métricas dirigidas hacia la instrucción de la intención de diseño, ya que la intención de diseño transferida a través de modelos CAD puede realizarse en tres etapas (croquis, datums de referencia y operaciones de modelado), con criterios contradictorios que deben ser equilibrados para llegar a la mejor estrategia de modelado.

La investigación ha incluido el desarrollo de un método de validación que demuestra que las rúbricas son dispositivos útiles para garantizar una comunicación consistente de la intención de diseño, y son fundamentales no sólo para evaluar, sino también para comunicar las expectativas del instructor. En esta investigación se examinó cómo definir claramente las cualidades de la intención de diseño (entre otras dimensiones de la rúbrica) para permitir una más fácil evaluación de un ensamblaje CAD.

Se ha mantenido la consistencia entre las rúbricas de piezas y ensamblajes, pero existen variaciones entre los descriptores y los niveles de desempeño para cada criterio, los cuales se redefinieron para ser más aplicables a ensamblajes. Fue necesario usar descriptores ambiguos y explícitos para evitar una valoración indiscriminada y asegurar una evaluación adecuada. Los niveles de desempeño reflejaron el nivel de conformidad para cada aspecto evaluado, con casos simples determinados dicotómicamente. Las dimensiones de la rúbrica se dividieron en seis categorías: válido, completo, consistente, conciso, claro y con intención de diseño. La plataforma Annota e-Rubrics se utilizó para los experimentos.

Para todas las dimensiones de la rúbrica, se encontró más concordancia y correlación entre instructores que entre instructores y estudiantes. Existe una correlación moderada/fuerte entre los instructores para las dimensiones de la validez, completitud, concisión y claridad, mientras que existe una ligera correlación para las dimensiones de consistencia e intención del diseño. Las pequeñas diferencias

entre los instructores permiten concluir que la rúbrica de ensamblajes propuesta proporciona una evaluación acumulativa objetiva de los estudiantes, por lo que se puede decir que los evaluadores pueden intercambiarse de forma segura.

Las contribuciones previas del equipo de investigación también fueron consideradas y mejoradas para esta tesis. En primer lugar, la rúbrica generó puntuaciones, con el proceso de puntuación definido por fórmulas, proporcionando una puntuación agregada de los niveles de desempeño. La rúbrica se mejoró usando tres características novedosas: dicotomización de criterios, criterios de evaluación con diferentes niveles de importancia y criterios "Pasa / No pasa".

En segundo lugar, las rúbricas también pueden ser descritas como estáticas o dinámicas. Las rúbricas estáticas, existen sólo en papel, no proporcionan retroalimentación inmediata al educando. Las rúbricas dinámicas realizan cálculos que proporcionan observaciones de evaluación inmediatas al usuario. Además, pueden adaptarse a situaciones específicas dependiendo de la capacidad del usuario. Las rúbricas electrónicas (eRubrics) son ideales para rúbricas dinámicas, y permiten el uso y desarrollo de rúbricas adaptativas y adaptables, como se describe a continuación.

En tercer lugar, las rúbricas deben ser adaptables (modificadas por el usuario), lo que debería hacerlas fácilmente comprensibles y fáciles de usar, y adaptativas (la rúbrica puede cambiar por sí misma, dependiendo del patrón de uso). Las rúbricas de evaluación se usan cuando un experto determina el progreso pedagógico de un educando, mientras que las rúbricas formativas son empleadas por los propios estudiantes, para trazar su progreso e identificar las deficiencias escolares para las que necesitan apoyo. Las rúbricas se deben refinar y mejorar de forma continuada, en un proceso iterativo y colaborativo, hasta que se alcance un acuerdo satisfactorio, tanto entre evaluadores como entre evaluadores y alumnos. Por ello, se desarrollaron mapas de aserciones que ilustran el modo en que la estrategia de expansión-contracción adapta las rúbricas al progreso del aprendiz de CAD, a la vez que ayudan a comprender las diferentes dimensiones de la rúbrica. Se desarrollaron materiales curriculares detallados de las dimensiones de la rúbrica (que constan de listas de control, buenas prácticas y diversas herramientas de evaluación) para investigar si son ventajosos para la comprensión del estudiante, especialmente cuando se necesitan compromisos para lograr diseños válidos.

Basándose en los experimentos con las rúbricas de ensamblajes, es evidente que las pequeñas diferencias entre los instructores sugieren que la rúbrica de ensamblajes propuesta es lo suficientemente sofisticada como para proporcionar una evaluación acumulativa imparcial del desempeño del alumno. En consecuencia, se puede afirmar con confianza que los evaluadores pueden usarse indistintamente sin sacrificar la precisión. Sin embargo, la rúbrica de ensamblaje posee una eficacia finita para producir una autoevaluación formativa de las habilidades de ensamblaje CAD para nuevos alumnos.

De todas las dimensiones de la rúbrica, el significado de la intención de diseño ha demostrado ser el más difícil no sólo de definir, sino también de transmitir a los estudiantes. Hay muchas razones para que esto sea así, pero sobre todo, la intención del diseño requiere recurrentemente un conocimiento previo preciso de cómo deberá funcionar el mecanismo, lo cual puede estar más allá del nivel de comprensión de estudiantes sin experiencia.

RESUM

Aquesta tesi posa el focus en el plantejament de fer servir rúbriques per a transitar de usuaris CAD bàsics a experts, seguint un enfoc pedagògic dissenyat per a produir models CAD d'alta qualitat. El mètode força un enfoc estratègic, donat que el coneixement estratègic (al contrari que el coneixement procedimental) està relacionat amb la identificació de procediments de treball alternatius i com fer la millor elecció entre ells per a garantir l'ús mes eficient de l'eina CAD.

L'avaluació del rendiment de l'estudiant es un component crític necessari per al aprenentatge dels estudiants implicats. L'ús de rúbriques per a dur a terme l'avaluació serveix no només com a un mètode per a que els instructors jutgin objectivament el treball dels estudiants, però també facilita l'autoavaluació del propi educand.

Aquesta investigació examinà l'ús de rúbriques de acoblament, descrivint la seua evolució a partir de rúbriques de peces, i estudià cóm afecten a la autoavaluació dels estudiants. També es va valorar la avaluació dels estudiants per els instructors, trobant que, mentre que les rúbriques de acoblament van ser parcialment compreses i fetes servir de manera eficient per els estudiants, van ser usades amb mes èxit per els instructors.

Es va fer una revisió exhaustiva de cóm ha estat definida històricament la qualitat del model CAD, molt sovint descrita mitjançant l'examen de la intenció de disseny i emfatitzant concurrentment la seua relació històrica amb la motivació de disseny. La motivació de disseny ha estat ben establerta en entorns industrials (en el context dels sistemes MCAD), mentre que la intenció de disseny segueix sent un concepte freqüentment mal entès. La intenció de disseny es defineix sovint com el comportament anticipat del model quant es sotmet a alteracions, i els mètodes de modelatge influeixen en la comunicació de la intenció de disseny, amb avantatjós significatius per a expressar la intenció del disseny mitjançant estratègies de modelatge adequades, especialment per a principiants CAD.

En aquesta investigació s'han abordat estrategues dissenyades per a millorar la comunicació de la intenció de disseny en models CAD, creixentat així la seua qualitat, amb directrius dirigides a avaluar la seua eficiència. Es evident que es necessiten mètriques dirigides cap a la instrucció de la intenció de disseny, ja que la intenció de disseny transferida a través de models CAD pot realitzar-se en tres etapes (croquis, datums de referencia i operacions de modelatge), amb criteris contradictoris que deuen ser equilibrats per a arribar a la millor estratègia de modelatge.

La investigació ha inclòs el desenvolupament de un mètode de validació que demostra que las rúbriques son dispositius útils per a garantir una comunicació consistent de la intenció de disseny, i son fonamentals no només per a avaluar, però també per a comunicar les expectatives de l'instructor. En aquesta investigació s'examinà com definir clarament les qualitats de la intenció de disseny (entre altres dimensions de la rúbrica) per a permetre una mes fàcil avaluació de un acoblament CAD.

S'ha mantingut la consistència entre les rúbriques de peces i acoblaments, però existeixen variacions entre els descriptors i els nivells de acompliment per a cada criteri, els quals es redefiniren per a ser mes aplicables a acoblaments. Va ser necessari fer servir descriptors ambigus y explícits per a evitar una valoració indiscriminada i assegurar una avaluació adequada. Els nivells de acompliment reflectiren el nivell de conformitat per a cada aspecte avaluat, amb cassos simples determinats dicotòmicament. Les dimensions de la rúbrica es dividiren en sis categories: vàlid, complet, consistent, concís, clar i amb intenció de disseny. La plataforma Annota e-Rubrics es va utilitzar per als experiments.

Per a totes les dimensions de la rúbrica, es va trobar mes concordança i correlació entre instructors que no pas entre instructors i estudiants. Existeix una correlació moderada/fort entre els instructors per a les dimensions de la validesa, completesa, concisió i claredat, mentre que existeix una lleugera correlació per a les dimensions de consistència i intenció del disseny. Las xicotetes diferencies entre

els instructors permeten concloure que la rúbrica de acoblaments proposta proporciona una avaluació acumulativa objectiva dels estudiants, per el que es pot dir que els avaluadores poden intercanviar-se de forma segura.

Les contribucions prèvies de l'equip de recerca també van ser considerades i millorades per a aquesta tesi. En primer lloc, la rúbrica va generar puntuacions, amb el procés de puntuació definit mitjançant fórmules, proporcionant una puntuació agregada dels nivells de conformitat. La rúbrica es va millorar fet servir tres característiques noves: dicotomització de criteris, criteris de avaluació amb diferents nivells d'importància i criteris "Passa / No passa".

En segon lloc, les rúbriques també poden ser descrites com estàtiques o dinàmiques. Les rúbriques estàtiques, existeixen sòls en paper, no proporcionen retroalimentació immediata a l'educand. Les rúbriques dinàmiques realitzen càlculs que proporcionen observacions de avaluació immediates al usuari. A mes, poden adaptar-se a situacions específiques dependent de la capacitat de l'usuari. Les rúbriques electròniques (eRubrics) son ideals per a rúbriques dinàmiques, i permeten l'ús i desenvolupament de rúbriques adaptatives i adaptables, como es descriu a continuació.

En tercer lloc, les rúbriques deuen ser adaptables (modificades per l'usuari), el que deuria ferles fàcilment comprensibles i fàcils d'usar, i adaptatives (la rúbrica pot canviar per sí mateixa, depenent del patró d'us). Les rúbriques d'avaluació se usen quant un expert determina el progrés pedagògic de un educand, mentre que les rúbriques formatives son fetes servir per els propis estudiants, per a traçar el seu progrés i identificar les deficiències escolars para a les que necessiten suport. Les rúbriques es deuen refinar i millorar de forma continuada, en un procés iteratiu i col·laboratori, fins que se arriba a un acord satisfactori, tant entre avaluadors como entre avaluadors i alumnes. Por això, es desenvoluparen mapes de assercions que il·lustren el mode en que la estratègia de expansió-contracció adapta les rúbriques al progres del aprenent de CAD, a la vegada que ajuden a comprendre les diferents dimensions de la rúbrica. Es desenvoluparen materials curriculars detallats de les dimensions de la rúbrica (que consta de llistes de control, bones pràctiques i diverses eines d'avaluació) per a investigar si son avantatjoses per a la comprensió de l'estudiant, especialment quant es necessiten compromisos per a assolir dissenys vàlids.

Basant-se en els experiments amb les rúbriques de acoblaments, es evident que les xicotetes diferencies entre els instructors suggereixen que la rúbrica de acoblaments proposta es lo suficientment sofisticada com per a proporcionar una avaluació acumulativa imparcial del acompliment de l'alumne. En conseqüència, es pot afirmar amb confiança que els avaluadors poden usar-se indistintament sense sacrificar la precisió. No obstant, la rúbrica de acoblaments posseeix una eficàcia finita per a produir una autoavaluació formativa de les habilitats de acoblament CAD per a nous alumnes.

De totes les dimensions de la rúbrica, el significat de la intenció de disseny ha demostrat ser el mes difícil no només de definir, però també de transmetre als estudiants. Hi ha moltes raons per a que açò sigui així, però per damunt de tot, la intenció del disseny requereix recurrentment un coneixement previ precís de cóm deurà funcionar el mecanisme, la qual cossa pot estar mes enllà del nivell de comprensió de estudiants sense experiència.

CONTENTS

Chapter I: Introduction 18

1. Introduction 18
2. Background 19
3. Goals and Objectives 21
4. Research Methodology 22
5. Contribution to Knowledge 22
6. Dissertation Outline 22

Chapter II: Systematic Review 24

1. Introduction 24
2. Design Rationale 25
3. Design Intent 27
   - 3.1. Representation Structures of Design Intent 30
   - 3.2. Capture and Transfer of Design Intent 33
   - 3.3. Design Intent Instruction 34
4. Conclusions 36

Chapter III: The Process for Assertions Maps and Rubrics Definition 37

1. Introduction 37
2. The Approach to Develop Rubrics 39
3. CAD Quality and its Dimensions: Related Work and Model Proposal 41
4. Quality Oriented Training 45
5. Assertions Maps 46
   - 5.1. Detailed Assertions 46
   - 5.2. Assertions Maps 48
6. More Lessons Learned on Using Rubrics 51
   - 6.1. Interspersed Rubrics 51
   - 6.2. Expand-Contract Process 54
   - 6.3. Teaching and Lecture Notes to Introduce Strategies and Tools to Evaluate Rubrics 60
7. Conclusions 65

Chapter IV: The Implementation of Part Rubrics 67

1. Introduction 67
2. Related Work 69

3. Specifications of an Adaptable Spreadsheet Rubric 70
4. Prototype of Adaptable Spreadsheet Rubrics 70
5. Experiments 72
5.1. Experiment 1 72
5.2. Other Results for Experiment 1 73
5.3. Experiment 2 75
5.4. Other Results for Experiment 2 76
5.5. Experiment 3 77
5.6. Other Results for Experiment 3 78
6. Discussion 79
7. Preliminary Conclusions 81
8. e-Rubrics Overview 81
9. Supplementary Background 82
10. Electronic Platform for Adaptable Rubrics 83
10.1. Design Specifications 83
10.2. Implementation 84
11. Experiments 86
11.1. Experiment 1 86
11.2. Experiment 2 89
11.3. Experiment 3 92
12. Conclusions and Future Developments 98

Chapter V: The Implementation of Assembly Rubrics 99
1. Introduction 99
2. Assembly Rubric Development 99
2.1. Descriptors and Achievement Levels 99
2.2. Rubric Scoring 100
3. Rubric Dimensions 101
3.1. Validity 101
3.2. Completeness 102
3.3. Consistency 102
3.4. Conciseness 102
3.5. Clarity 103
3.6. Design Intent 103
3.7. Summary 104
4. Rubric Rating Scale 106

5. Experiments 107
5.1. Experiment 1 107
5.2. Experiment 1 Discussion 109
5.3. Experiment 2 117
5.4. Experiment 2 Discussion 118
6. Conclusions and Recommendations for Future Work 127

Chapter VI: Conclusions and Future Work 129
1. Summary of Findings 129
1.1. Objective 1: Literature Review 129
1.2. Objective 2: Rubric Development 130
1.3. Objective 3: Rubric Validation 131
1.4. Objective 4: Development of Pedagogical Tools 132
2. Conclusions 132
2.1. Rubrics 132
3. Future Work 133
3.1. Rubric Refinement 133
3.2. Rubric Tools 134
3.3. Automating Rubrics 134
3.4. Rubrics in Education 134
3.5. Drawings Rubrics 134
3.6. Rubrics for Hybrid Surface and Direct Modeling 135
4. Concluding Remarks 135

REFERENCES 136
APPENDIX A: Assembly Rubric Presentations 147
APPENDIX B: Statistics Results 166
APPENDIX C: Table 3.7 Enlarged 207

LIST OF FIGURES

Figure 2.1. Schema illustrating current dispersion in understanding of design rationale 26
Figure 2.2. Schema illustrating current dispersion in understanding of design intent 28
Figure 2.3. Proper labeling of modeling operations as a simple way to convey design intent 30
Figure 2.4. Three approaches to model a casing 31
Figure 2.5. Example of reference datum supported by one line embedded in a profile 32
Figure 2.6. Addition of a second leg using symmetry 32
Figure 3.1. Casing with flaps 52
Figure 3.2. Shut-off valve 52
Figure 3.3. Full rubric with interspersed assertion on modeling, assembling and extracting drawings 53
Figure 3.4. Plughole assembly 55
Figure 3.5. Lower plughole body 56
Figure 3.6. Pulley 59
Figure 3.7. Profiles with different constraints result in different flexibility 61
Figure 3.8. Profiles are flexible if shape, size, position and orientation are fixed independently 61
Figure 3.9. Pick and try to drag different lines and nodes to visualize unconstrained degrees of freedom 62
Figure 3.10. Complex profile to draw and self-evaluate with rubrics during practice 1 63
Figure 3.11. Scoring rubric of practice 1 63
Figure 4.1. Front end of the prototype adaptable rubric showing the combo box for criterion 1 71
Figure 4.2 Front end of the prototype adaptable rubric with radio buttons 71
Figure 4.3. Illustration of how level 2 is shown for level 1 first criterion 71
Figure 4.4. Assemblies used for modeling in experiment 1 72
Figure 4.5. Detailed drawing of the Bertoline et al. part to model during experiment 2 75
Figure 4.6. Detailed drawing of the part to model during experiment 3 77
Figure 4.7. Annota framework 84
Figure 4.8. Assembly rubric as available in the Annota platform 85
Figure 4.9. Annota rubrics are adaptable by allowing users to fold and unfold levels of detail 85
Figure 4.10. Criterion with an associated scoring weight 86
Figure 4.11. Example of bubble with textual explanation for a particular level of deployment 86
Figure 4.12. Detailed drawing of the part used in experiment 1 87
Figure 4.13. Pneumatic cylinders used in experiment 1 88
Figure 4.14. Detailed drawing of the part used in experiment 2 90
Figure 4.15. Axonometric view of the pistol clamp used in experiment 2 90
Figure 4.16. Parts used by experimental group and control group in experiment 3 93

Figure 4.17. Assemblies used by the experimental group and control group 96
Figure 5.1. Sample assembly rubric illustrating performance levels 100
Figure 5.2. Non-standard parts used for modeling in experiment 108
Figure 5.3. Sub-assemblies used in experiment 1 108
Figure 5.4. Final assembly in experiment 1 109
Figure 5.5. Sectioned filter assembly used in experiment 2 118

LIST OF TABLES

Table 3.1. Assertions for modeling 47
Table 3.2. Assertions map for models 49
Table 3.3. Differences between the self-evaluation and the teacher evaluation in experiment 1, task A 54
Table 3.4. Differences between the self-evaluation and the teacher evaluation in experiment 1, tasks B and C 54
Table 3.5. Rubrics of undergraduate students 57
Table 3.6. Rubrics of Master's students 57
Table 3.7. Condensed rubrics of the six parts of the pulley 60
Table 3.8. Differences between the self-evaluation and the teacher evaluation in experiment 3 64
Table 3.9. Teacher evaluation in experiment 3 65
Table 4.1. Comparison between instructor, self and peer evaluations for experiment 1 74
Table 4.2. Comparison between the instructor, self and peer evaluations for experiment 2 76
Table 4.3. Comparison between instructor, self and peer evaluations for experiment 3 79
Table 4.4. Percentages of agreement between self-evaluations and instructor evaluation for modeling task 88
Table 4.5. Percentages of agreement between self-evaluations and instructor evaluation for assembly task 89
Table 4.6. Percentages of agreement between self-evaluations and instructor evaluation for drawing task 89
Table 4.7. Percentages of agreement between self-evaluations and instructor evaluation for experiment 2 91
Table 4.8. Percentages of agreement between the students using bubbles and the rest 91
Table 4.9. Time each detailed explanation was displayed for the modeling rubric in experiment 2 92
Table 4.10. Percentages of agreement in the control group (0.25) 94
Table 4.11.Percentages of agreement in the experimental group (0.25) 94
Table 4.12. Percentages of agreement between peer students and instructor (0.25) 95
Table 4.13. Percentages of agreement between instructors (0.25) 95
Table 4.14. Percentages of agreement in the control group (0.25) 97
Table 4.15. Percentages of agreement in the control group (0.05) 97
Table 4.16. Percentages of agreement in the experimental group (0.05) 97
Table 5.1. Validity scores for mid-term exam 110
Table 5.2. Completeness scores for mid-term exam 111
Table 5.3. Consistency scores for mid-term exam 112
Table 5.4. Conciseness scores for mid-term exam 113

Table 5.5. Clarity scores for mid-term exam 114
Table 5.6. Design Intent scores for mid-term exam 115
Table 5.7. Inter-rater reliability scores for mid-term exam 116
Table 5.8. Pearson correlation values for mid-term exam 117
Table 5.9. Mid-term exam p-values for Wilcoxon Signed Rank test 117
Table 5.10. Validity scores for final exam 120
Table 5.11. Completeness scores for final exam 121
Table 5.12. Consistency scores for final exam 122
Table 5.13. Conciseness scores for final exam 123
Table 5.14. Clarity scores for final exam 124
Table 5.15. Design Intent scores for final exam 125
Table 5.16. Inter-rater reliability scores for final exam 126
Table 5.17. Pearson correlation values for final exam 126
Table 5.18. Final exam p-values for Wilcoxon Signed Rank test 126
Table 5.19. Percent agreement values between mid-term and final exam 127
Table 5.20. Pearson correlation values between mid-term and final exam 127

Chapter 1
Introduction

1. Introduction

CAD model quality [Contero et al., 2002] has historically been introduced and enforced by following style guides and standards, or more recently, by using tools provided by modeling software, but to date, only low semantic level aspects of CAD model quality have been principally considered. Early introduction of quality model construction to novice learners alleviates future problems in their development when tasks and models become more complex. The vision for this research was to design and evaluate CAD modeling rubrics (specifically for models and assemblies) in order to validate their effectiveness in assisting novice CAD users to design quality models. Protocols for student use of CAD modeling tools in order to assess model quality are also analyzed, but increasing the semantic level of CAD model quality was another objective. Toward this end, the final goal is to expand the definition of design intent to incorporate traditional understanding of design rationale for all parametric models.

Rubrics have the potential to serve as vital tools used to enforce quality modeling techniques and adherence to industry standards for novice CAD trainees. Research and rubric development for parametric part files has previously been established, although additional testing continues. A logical outcome in the continuation of this process is the design and development of rubrics for corresponding assembly files used to support product data quality and collaborative engineering throughout the design cycle, while also quantifying design intent in these models leading to the development of a taxonomy of preferred practices.

In order to avoid confusion, for the purposes of this thesis, the acronyms CAD and MCAD are used interchangeably. While these terms have distinct definitions, an attempt has been made to be consistent with the designations used in the cited literature. While the focus of this research is mechanical and manufacturing in origin, there exist underlying similarities in CAD processes and approaches that are common in multiple areas, which could be beneficial to those specific pursuits. Furthermore, it should be understood that this research is concerned primarily with feature-based modeling.

2. Background

Two separate problems have been identified: (1) CAD quality concepts are introduced late in novice CAD learner development, and (2) CAD quality concepts are poorly understood, so the topic of quality needs to be further investigated so that enhanced strategies in dealing with higher semantic levels of quality are obtained.

CAD instruction should not emphasize CAD quality as a complementary goal to be addressed only after basic skills have been obtained. Instead, CAD quality should be the primary focus of the training process from the very beginning of instruction. It is true that many beneficial practices aimed at increasing the quality of CAD models have been used to promote new teaching strategies (i.e. [Company et al., 2007]), but they have not yet been aligned in a comprehensive approach that involves both students and teachers.

Using the classification by Field, the intent is to narrow the gap between the majority, the experts, and the super-users by improving the training period of the majority so they can attain an expert level [2004].

Rubrics have been successfully examined as a valid means for disclosing quality criteria that can be efficiently transmitted by way of good practices in a basic CAD course for fresh designers/engineers [Company et al., 2013]. However, further tests appear to conclude that rubrics are useless if they are generic and are used in a standard manner. Only detailed rubrics, used under suitably designed experimental conditions, are effective in disclosing which specific strategies to enforce quality may be understood early in the training process, and how they can be better introduced.

International Organization for Standardization (ISO) 9000:2005 defines quality as the degree to which a set of inherent properties of a product or process fulfills implicit or explicit needs [2005] while Freiesleben describes design quality as the degree to which products meet customer needs [2010]. Although the concept of quality in mechanical CAD models has been widely studied and includes numerous sets of peculiarities and open problems, a wide-ranging and definitive overview was described by Contero and colleagues [2002].

Two key inter-related concepts within the product design discipline are Design Rationale and Design Intent. A recent contribution in this area is due to research conducted by Zhang and collaborators, where they emphasized the relationship between design intent and design rationale, while also investigating why relatively few design rationale systems have been practically implemented in industry [2013]. Karsenty provided a significant contribution by measuring goodness of captured design rationale [1996], while Bracewell and colleagues described a strategy to implement tools used to capture, represent, and retrieve design rationale [2009].

Lee and Lai developed a framework to acquire and assess design rationale representations to discern explicit elements of design rationale while supporting multiple tasks [1991]. Mostow states that any model of the design process should communicate the state of the design, the goals of the design process, design justifications, and the role of learning [1985]. Furthermore, the benefits of design rationale representation include communication and helping to delineate different levels of abstractions [Mostow, 1985], [Fischer et al., 1991], and [MacLean et al., 1991].

Research by Johnson and Diwakaran provides insight into design intent, as they disclose that creating designs too quickly adversely affect design intent, while model quality should correlate with the time required for model revision [2009]. In continuation of their research, it was found that using simpler features increased the time required for artifact creation, but increased model reuse [Diwakaran and Johnson, 2012]. Plumed and colleagues investigated methods to determine design intent embedded in 2D sketches by dissecting the drawing into features and subsequently analyzing these combinations of features, which can then be catalogued and identified [2012].

Rynne and Gaughran stated that CAD software records the succession of features used to create the model, thereby reflecting the user's opinion of the best approach [2008]. Zhang and Luo state that

CAD illustrates design intent through an artifact's history, features, parameters, and constraints, and explored methods to share design intent between models, but encountered difficulties resulting from the absence of standards in data-exchange procedures [2009].

Protocols for student use of CAD modeling tools used to assess model quality can be found in standards such as ISO 9000:2005. To this end, it must be noted that some texts recompile most relevant standards as evidenced by Lieblich and Fischer [2008]. Finally, Johnson and Diwakaran provide a recent example of an approach directly aimed at student education [2011].

It must be recognized that the communication of complex design information is frequently lost between stages in the design cycle. Numerous CAD software incorrectly define precise areas on an artifact as a "feature," when they are more accurately defined as modeling operations, or "form" features. Three types of features need to be acknowledged: design features (features corresponding to model function) [Vandenbrande and Requicha, 1993], manufacturing features (shapes or areas created by various manufacturing processes), and form features (geometric shapes with no corresponding relationship with model function) [Han, 1996], [Han, 1997]. In fact, manufacturing processes can change and improve over time while the artifact fulfills the same function [Vandenbrande and Requicha, 1993]. As such, design intent is more difficult to convey through form features than design features.

In summary, while CAD software can record the sequence of steps used to create the artifact, it cannot relate why certain commands were used. Methods need to be developed so that this information can be documented and design justifications understood. As a further aside, it would be highly beneficial if the extraction of this information could be represented in graphical format.

In order to further clarify the main dimensions of CAD model quality, the following dimensions have been proposed by Company and colleagues [2013]:

1. Models are valid if they can be successfully accessed by suitable software which result in no errors or warnings.
2. Models are complete if all necessary product characteristics are provided for all design purposes.
3. Models are consistent if they do not crash during normal design exploration or during common editing.
4. Models are concise if they do not contain any extraneous (repetitive or fragmented) information or techniques.
5. Models are clear and coherent if they are understood at first glance.
6. Models are effective if they convey design intent.

Conversely, in reality, CAD software does not literally manipulate physical objects. They only control data which represents these objects [Requicha, 1980]. As such, designers have no control over how the information is stored and visualized. What is essential is how the models convey the required information, which is in this case design intent. In fact, "Dimension 5" (Clarity) and "Dimension 6" (Effective Conveyance of Design Intent) were in reverse order, but were interchanged because "Dimension 6" was found to be more complex as it concerns higher semantic level concepts.

General concepts about rubrics can be easily reviewed in research conducted by Goodrich, where she states that not only are rubrics assessment tools, but they also specify gradations between quality levels [1996]. These assessments at incremental levels are vital for increasingly complex tasks introduced throughout the instructional curriculum.

Another common strategy aimed at increasing quality relies on the development of software tools in order to test (and occasionally automatically repair) CAD models. These strategies are a small subset of what Booker classifies as "design tools" [2012] . Typical examples include specific tools, oftentimes embedded in CAD applications, used to detect short lines or isolated vertices, and have

been named "CAD Model Quality Testing (MQT) tools" by González-Lluch and colleagues [2017a and 2017b].

Additionally, preceding research has been established in introducing new teaching strategies [Company et al., 2007], working on CAD quality concepts [Contero et al., 2002] , and even in using rubrics to convey quality criteria [Company et al., 2013].

3. Goals and Objectives

Three main reasons are foreseen as to why CAD practice fails in enforcing quality. First, careless modeling strategies result in errors not always detected by the CAD application. Second, even best practices aim at the very modest goal of pure geometric quality. High level best practices, like those aimed at emphasizing design intent, are rarely reinforced at every stage, nor transmitted between stages. Finally, the lack of quality in CAD instruction can also be blamed for failure to reinforce quality.

In order to facilitate these steps forward, the following objectives will be addressed:

1. Bibliography search and analysis to review the concept of CAD model quality in order to determine how the dimensions of quality models can be detailed in suitable assertions and how good practices may be organized around these main dimensions in a hierarchical way that allows the attainment of a new set of rubrics, developed to explicitly make quality the main goal for both the teacher and the students. The hierarchical structure of this new set of rubrics is designed to allow varying levels of detail, which align with the different stages in the evolution of the learning process. This process allows adaptable evaluation which prevents model quality from being only evaluated at the end of instruction.

   The bibliographical search includes analyzing style guides from different original equipment manufacturers (OEM) and standards (such as ASME Y14.41-2012 or ISO 16792-2015).

2. An original set of rubrics are created and designed (with associated graphical assertions maps), in order to not only emphasize proper modeling practices, but also to stress the importance of creating robust models through proper design intent so that the corresponding assemblies can be updated with minimal effort. These rubrics are hierarchical in nature, with increasing quality dimensions added as the task sequence becomes more advanced. Furthermore, these rubrics provide metrics for objective measurement of incremental quality linked to good practices.

3. These rubrics are tested and statistically validated in a classroom setting so that objective determinations can be made about the link between a specific procedure and the quality of the corresponding model created. Qualities verified include validity (the model can be opened and examined without error), succinctness (the model has the exact amount of information needed to define it and nothing more) robustness (the model can be altered and updated without failure, even between file types), and design intent quantification (assessment of model's creation). These qualities are not only validated (and specifically defined in order to allow quality quantification) within each specific file, but also examine the relationships between all interlinked files (parts and assemblies). Specific definitions of these qualities are proposed to facilitate fundamental assessment.

4. As numerous current parametric modeling software already contain tools that assist in testing CAD model quality and repairing poorly designed models, strategies have been developed to increase the utilization of these tools, especially in the early stages of trainee instruction. Early introduction of these techniques is useful so that proper practices can be instilled (and negative practices eliminated) so that poor part creation does not convey complications with

the corresponding assembly files. Detailed protocols are developed and tested so that CAD learners can self-test and self-assess their models for specific quality criteria.

4. Research Methodology

As with all research, a hypothesis has been provided, with supporting data collection and analysis theoretically used to support this premise. Various qualitative rubrics are developed and tested. Embodiment design, as detailed in Pahl and collaborators, is utilized in order to identify specific rules and processes used in various CAD models and their file dependencies [2007].

Specifically, several experiments are conducted throughout this thesis, examining the use of rubrics in constructing part and assembly models. In some experiments, a control and experimental group were utilized, when the research aim was examination of improved modeling performance. In other experiments, most notably assembly modeling, the purpose was to detect if the rubrics were well-understood between students and instructors. In such cases inter-rater reliability, Pearson Correlation, and Wilcoxon Signed Rank Tests were investigated.

5. Contribution to Knowledge

Benefits of this research include:

1. Further the clearly defined qualities of design intent (and model quality) in such a manner that lends itself to easy assessment.
2. Construction of assessment rubrics to accurately represent comprehensive model quality and design intent depiction.
3. Standardization of these definitions and assessment strategies.
4. Development of detailed protocols for learners so that they can self-assess whether their models achieve expected quality criteria.

6. Dissertation Outline

This dissertation is divided into the following chapters.

Chapter 1: Introduction

Introduction to the concept of design intent and rationale, the definitions of CAD model quality, and the need for rubrics to chart student progress with the provision for self-assessment. Benefits of this research, along with the contributions to new knowledge are provided.

Chapter 2: Literature Review/State-of-the-Art

A review of historical research is provided, along with a discussion of design intent, design rationale, and the role of rubrics to convey CAD quality model concepts. A discussion is provided, focusing on existing gaps in CAD quality modeling rubrics and how they can be adapted to assembly modeling.

Chapter 3: Adaptable Rubrics

Rubrics increase a novice's ownership of their learning and decreases the amount of time needed for instructor assessment [Goodrich, 1996], but rubrics must accomplish different specifications. In order to distinguish user-driven rubrics from system-driven rubrics, the terms "adaptable" and "adaptive" are adopted. While an adaptable rubric can be modified by the user to adapt to different needs, an adaptive rubric is able to adapt or change itself, depending on the usage pattern (ex. software that deploys a rubric's level of detail depending on analysis of a student's modeling procedure). A discussion of rubrics is furnished detailing

how they must be not only adaptable, but also adaptive. Adaptable rubrics should also be *practicable* (easily understood and user-friendly).

Chapter 4: Implementation of Rubrics

This chapter addresses the idea that rubrics provide feedback but also evolve dynamically so that differing learning paces are accomodated. Computer-Assisted Assessment (CAA) tools can be used to provide automatic feedback, but are often not customizable, thereby preventing adoption. The need for adaptable and also practicable rubrics arises and a discussion of implementation strategies and tools for adaptable rubrics is provided.

Chapter 5: New Rubrics and Validation

This chapter discusses the creation of an original rubric, which illustrates that the methodology and tools described in this thesis are generic enough to serve as the basis for the development of new rubrics. This chapter further provides a detailed explanation of the experiments used to validate the novel rubric developed with Annota eRubrics, a web-based structure used to manage adaptable rubrics and students. It is illustrated, that using Annota, rubrics can be managed and results can be accessed while also registerig new students.

Chapter 6: Conclusions and Future Work

This chapter includes a summary of results and includes conclusions and ideas of future developments. Design intent and design rationale are the root concepts to convey highly semantic quality criteria in CAD modeling through good practices. While rubrics are supportive for this purpose, for rubrics to convey complex and multidimensional goals, they must be both adaptable and practicable. Novel tools are required to deal with adaptable rubrics and the tool “Annota” has been proven valid for this purpose.

Introduction of complex rubrics in educational sceanarios must be bottom-up rather than top-down. Rubrics and introduced top-down when the main dimensions are introduced first (validity, completeness, consistency, conciseness, clarity, and conveyance of design intent). If they are presented bottom-up, the detailed criteria to check those concepts is introduced without explicitly explaining the general concept to which they belong. As an example, instruction may emphasize that sketches must be fully constrained without explicitly clarifying that this task is a measure of the robustness dimension. Furthermore, rubrics should only contain quantifiable criteria, and this implies developing new MQT tools since current tools are not acceptable.

Chapter 2
Systematic Review

1. Introduction

Feature-based parametric CAD is a commonly deployed 3D modeling technology that is widely used in industrial settings. In these systems, the 3D CAD model is created by gradually and sequentially adding geometric features through parent/child relationships, which creates an interconnected structure that, when defined properly, allows for more flexible and reusable models. This process is recorded in a structure known as a design tree, feature tree, or history tree (depending on the modeling package).

Parent/child interdependencies are the basic elements that facilitate CAD reusability and alteration of parametric models. When these dependencies are properly defined, changes in the artifact can be performed efficiently, as alterations propagate automatically from parent to child nodes. However, parent/child dependencies can also be the root of numerous regeneration problems, which often forces designers to recreate the CAD model entirely, costing time and money.

Previous researchers have determined that 48% of CAD models fail during design exploration [Jackson and Buxton, 2007] and according to the 2013 State of 3D Collaboration and Interoperability Report, 49% of engineers spend more than 4 hours per week repairing design data, with 14% spending more than 24 hours per week [Jackson and Prawel, 2013]. The same report states that 32% of organizations miss deadlines due to design data problems [Jackson and Prawel, 2013]. Gerbino states that data exchange issues result from poor modeling strategies [2003]. González-Lluch and colleagues echo these sentiments, stating that erroneous CAD models that filter toward downstream applications require effort to rework the models to remove data corruption [2017a]. Poor understanding and/or communication of design rationale and design intent are commonly argued to cause most of those failures. But the concepts of design rationale and design intent are complex in and of themselves.

Describing the purpose of a design and the justifications for specific decisions made when creating it are essential tasks for engineers and design professionals. Design rationale can be defined as the explicit documentation of the reasons behind the decisions made when designing a system or artifact [MacLean et al., 1989]. Although design rationale applies to a number of disciplines [Regli et al., 2000], its representation and management significantly vary across diverse fields. For example, communication of design rationale through source code comments has been a relevant research topic in software engineering for a number of years [Jarczyk et al., 1992]. In product design however, special tools and approaches are required, as the artifacts used to represent design in CAD systems are much more complex than simple text comments. Despite the ongoing research, fixing a generally agreed definition of the term and finding efficient mechanisms to convey design rationale in product and engineering design are still open issues that cause a noticeable lack of general tools to support design rationale communication.

The chapter compares existing definitions of design rationale, determining that design intent communication is a crucial initial step towards understanding design rationale. Further, the chapter addresses the definition of design intent, concluding that a standardized manner in which to explicitly communicate or deduce a CAD model's design intent does not yet exist. Wang and colleagues, who studied a push system to provide shared design knowledge, support this view, recognizing that "no existing knowledge acquisition method is satisfied to support mechanical conceptual design [2015]."

The review of the current understanding of design intent and its historical connection to design rationale is presented, focusing on the difficulties encountered when conveying design intent through CAD model geometry, resulting in most parametric modeling applications offering various complex sets of tools to manage this information. In this context, it is recognized that current research [Company et al., 2015] focusing on defining quality metrics used to verify that design intent is properly incorporated into the modeling strategy is successfully implemented when constructing the CAD model. Thus, it is finally suggested that the capture and transfer of each type of design intent is manifested at its most appropriate representation level: sketches, modeling operations, relationships between modeling operations, and reference datums (which include axes, planes, etc. used to link the model to the view and/or components in the model. In summary, the current state of design intent instruction is addressed, with recommendations for future advances.

2. Design Rationale

Before the idea of design rationale became commonplace, industrial products and their components were described exclusively in terms of how they functioned, but not why they were designed in a certain way [Regli et al., 2000]. As a result, the time and communication effort required of collaborative teams to reason and understand each other's designs increased significantly as projects grew in complexity. Design rationale systems were introduced as basis of reasoning and communication among such teams [Chandrasegaran et al., 2013].

The term design rationale has historically been defined in a variety of ways. For example, Shum and Hammond defined it as "elements of the reasoning which has been invested behind the design of an artifact [1993]." Sim and Duffy describe it as "the reasoning and argument that leads to the final decision of how the design intent is achieved [1994]." The same authors also define design intent as the "expected behavior that the designer intended the design object should achieve to fulfill the required function [Sim and Duffy, 1994]." A more complete definition was suggested by Lee: "Design rationales include not only the reasons behind a design decision but also the justification for it, the other alternatives considered, the tradeoffs evaluated, and the argumentation that led to the decision [1997]."

Szykman and colleagues define design rationale as the documentation of the design intent of an artifact [2001]. They also contend that schemes are needed to retrieve, clarify, and facilitate exploitation of design information. These systems should capture and represent the progression of design intent, comprehension about the model throughout the development process, and associations, which link decisions. ISO defines design rationale as the, "logic underlying the methodology and used in constructing the design [ISO, 2005b]."

Design rationale consists of different types of information, such as the history of the design process and the reasons for making each decision. This knowledge can be useful at various stages such as design verification, evaluation, reuse, teaching, communication, documentation, and maintenance [Burge and Brown, 1998].

The state of the art for defining design rationale is summarized in the IBIS-like schema (Issue-Based Information Systems) shown in Figure 2.1. An IBIS schema is an approach to represent complex problems that involve multiple stakeholders. It was invented by Kunz and Rittel [1970] and is the base on top of which new schemas are being developed (such as ISAA—Integrated Issue, Solution, Artifact, and Argument—by Zhang et al. [2013]).

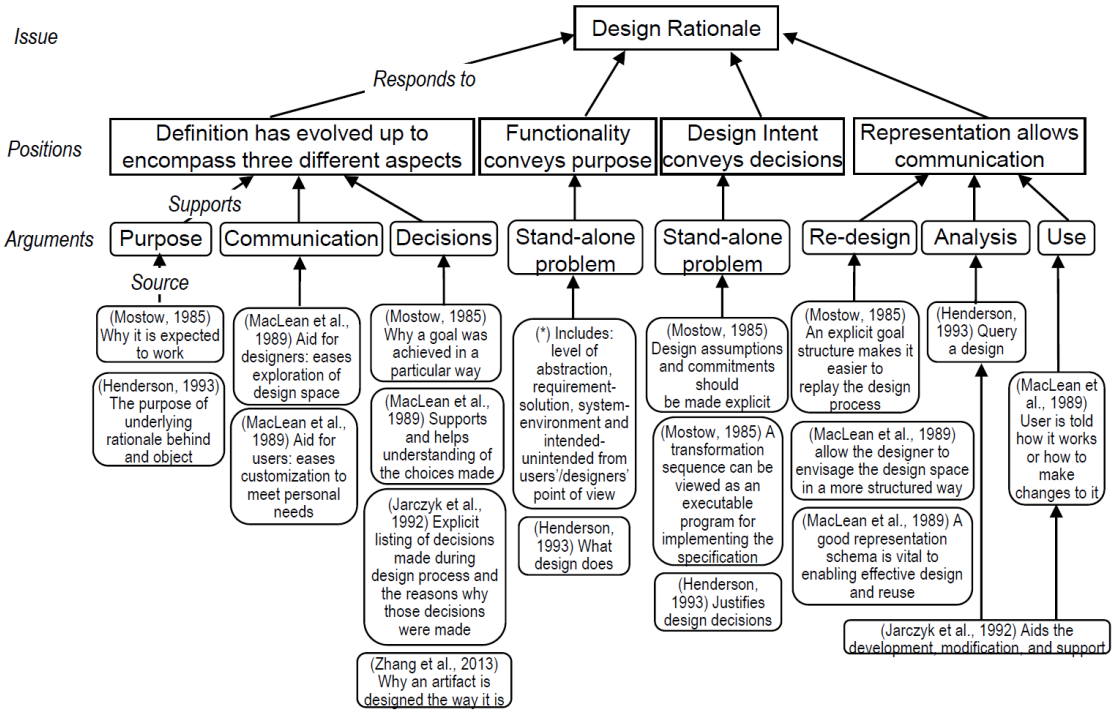


Figure 2.1. Schema illustrating current dispersion in understanding of design rationale.

As shown in the above schema, Mostow [1985] first realized the importance of making design rationale explicit, but his work was directed toward finding better models of the design process. In investigating the global design progression, he stated that design rationale is just one step in the design process. According to him, design rationale clarifies and justifies why a certain decision was made and why it was thought to be the correct path to take. Design rationales need to be both explicit (clearly defined goals) and appropriate (reasons given why a certain path was chosen) [Mostow, 1985].

MacLean et al. [1989] focused on defining and representing design rationale, highlighting its role as an aid for both designers and end users. The researchers emphasize its importance, describe its benefits, and develop a "semi-formal" notation to make it explicit [MacLean et al., 1991]. Unfortunately, their representation is aimed at computer software design and does not consider product design peculiarities. Lee and Lai [1991], also focusing on software design, highlighted the importance of selecting a suitable representation, and provided a framework for evaluating a design rationale representation. This framework increasingly discerns explicit elements of design rationale and supports multiple design tasks. They discuss and evaluate Decision Representation Language (DRL) in order to accomplish these tasks.

In attempting to integrate physical and conceptual models, Henderson divided product models into physical and meta-physical domains [1993]. The physical domain integrates all information related with a model's actual manifestation, such as geometry, dimensions, and materials while the meta-physical realm refers to information that describes the structure and behavior of the model. It is argued that metaphysical modeling provides the capability to capture the function and design intent of systems, assemblies, parts, features, and even individual dimensions and tolerances. This modeling process uses Product Definition Units (PDU), which are shells used to encapsulate information. Henderson indirectly defines design rationale, as he describes design intent as "the purpose or underlying rationale behind an object [1993]." While this definition does not represent the current understanding of design intent, the term attempts to explain the difference between intent and functionality ("intent justifies a design decision whereas the functionality just tells what the design does").

Karsenty evaluated the importance of representing design rationale in cases where the original design is reused [1996]. His research questioned six designers about their need to understand previous design rationale, how archived design rationale was utilized, and how to effectively acquire design rationale. He states that design rationale could be beneficial for those requiring reinforcement for design-based decisions, but it is not adequate to be used as the sole support. In fact, he used the QOC Notation originally developed by MacLean et al. [1989] to document design rationale.

In addition to the review of early contributions, open problems, and a classification of systems and tools for design rationale capture and retrieval in the context of CAD tools, the work by Regli et al. [2000] is also compelling as it clearly identifies the multidisciplinary nature of design rationale. According to the authors, a problem develops when design collaboration is needed and communication is absent; and design rationale is crucial to avoid these problems. Their work states that the need for design rationale is a collective problem, encountered in all industries, but design rationale systems are uncommon. Design rationale systems need to assess design approaches, representation schema, capture, and retrieval. A system, which could capture such information, would be important for those tasked with managing design data.

A recent study examined the understood purpose of a design between industrial designers and design engineers in order to increase collaboration [Laursen and Møller, 2015]. They found that complex and abstract industrial design elements (ex. meaning, emotions, etc.) were less shared with engineers and that various aspects of the design are perceived separately, ensuring that the shared knowledge is less meaningful to the engineers [Laursen and Møller, 2015].

Additionally, the extent to which we can benefit from design rationale depends largely on the language used to represent it [Lee and Lai, 1991]. The work by Karsenty [1996] is a significant contribution in regards to measuring the goodness of captured design rationale. The work by Bracewell and colleagues [2009] is also noteworthy, as it describes a strategy to implement customized tools to capture, represent, and retrieve design rationale.

A more recent contribution in the area of design rationale is due to Zhang and researchers [2013], which not only highlights the relationship between design intent and design rationale, but also investigates why only a small amount of design rationale systems have been implemented in industry. It appears that the limitations exhibited by the traditional approaches to capture design rationale summarized by Karsenty [1996] and recently addressed by Bracewell et al. [2009] are still valid.

As shown in Figure 2.1, design intent is a significant contributor to design rationale, but it can be studied as a stand-alone problem, which will be considered in the next section. Design rationale describes the purpose of a design, the reasons relating why certain steps were taken in artifact creation, and also aids communication in a collaborative environment, particularly for end users. Functionality conveys purpose, and the literature on function reveals that this is a separate ambit where there exist many views of function, and not all of these views are made explicit [Srinivasan et al., 2012]. It is concluded that accepting the multifaceted nature of design rationale is a mandatory aspect to confront the unsolved problem of finding a suitable language to represent it.

3. Design Intent

Design intent is a nebulous concept. Some authors have even stated that a formal definition of the term is problematic to obtain [Chen and Hoffman, 1995], although many have attempted to describe it. Others use the term without providing any definition [Ault, 1999a]. ISO defines design intent as the, "intentions of the designer of a model with regard to how it may be instantiated or modified [ISO, 2005c]." In a pioneering work by Requicha and Rossignac, they do not explicitly define design intent, but emphasize that product models contain unambiguous information about behavior and function and most relevant data is associated with design features [1992]. In reality, it is a common assumption that a standard definition is understood already, as many authors use the term "design intent" without providing an explicit definition while other researchers use implicit references to it. Wiebe states that

the use of CAD models implies that information traditionally documented in working drawings is now housed in the model database [1999].

In a survey conducted by Iyer and Mills [2006], common elements were identified in a number of definitions and interpretations of design intent. This information was used by them to provide a comprehensive definition in the domain of 2D CAD: "Design intent contained in legacy CAD is the insight into the design variables (design objectives, constraints, alternatives, evolution, guidelines, manufacturing instructions and standards) implicit in the structural, semantic and practical relationships between the geometric, material, dimensional and textual entities present in the CAD representation." [Iyer, and Mills, 2006]. They also acknowledged the capture, representation, and retrieval of design intent as open issues for future research.

Although a lack of consensus exists within the scientific and technical community on the exact definition of design intent, there is agreement on its importance and the benefits of an explicit representation. Advantages were summarized by Pena-Mora et al. [1993] in the form of four points:

- Changes in complex projects require certain design decisions to be modified during the development process. When the justifications defined during the initial stages are lost, they need to be recreated, which has a negative impact on project costs and development times. The ability to store, process, and retrieve this information can significantly improve productivity.
- When design intent information is represented explicitly and is easily available for review, the overall quality of the product increases.
- Explicit representation of design intent leads to a more intelligent use of resources and knowledge.
- Efficient communication of design intent is essential for integrating solutions and transferring design knowledge.

An IBIS-like schema summarizing the state of the art for defining design intent is illustrated in Figure 2.2.

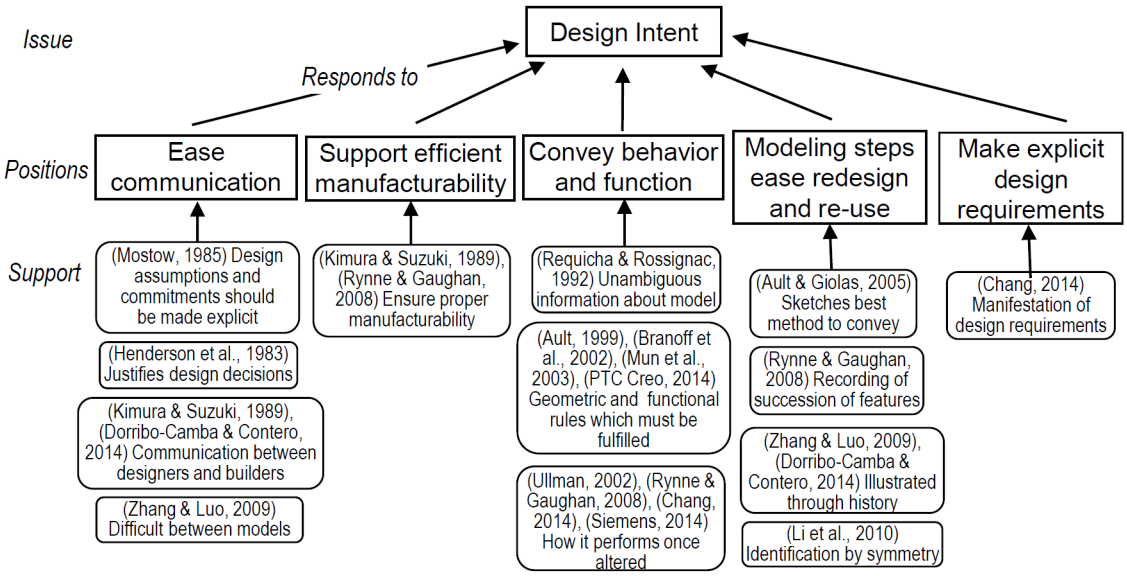


Figure 2.2. Schema illustrating current dispersion in understanding of design intent.

The schema reveals the multifaceted nature of design intent, as it mainly conveys behavior and function, but it also makes design requirements explicit and eases communication. In doing so, design intent facilitates redesign and reuse, and even efficiently supports manufacturability.

Mun et al. define design intent as the cluster of geometric and functional rules that must be fulfilled by the final product, represented by parameters, constraints, features, and history [2003]. Kimura and Suzuki define design intent as the way original designers articulate the objectives of the design so that the manufacturer can understand the design process in order to ensure proper manufacturability without hampering design performance [1989]. Design intent defined in this manner incorporates design requirements, behavior, and function while facilitating communication between designers and builders. They further state that design intent plays a vital role in communication in simultaneous design.

In their research on modeling strategies in CAD pedagogy, Rynne and Gaughran define design intent as a description of how an object is modeled and also how it should perform once it is altered [2008]. They also assert that CAD software records the succession of features used to create a model, which reflects the user's opinion of the best approach to accomplish a specific task. They further state that design intent should be more comprehensive than shapes and sizes of features, but must encompass consideration of manufacturing methods and relationships between features. A student's ability to accurately model an object correlates with their ability to visualize and assemble the objects cogently.

Ullman alludes to a consensus among the CAD community whereas intent exemplifies the arranging of geometric constraints in a parametric system. This classification defines the geometric dependency needed by the system in order to enable alterations [Ullman, 2002].

Ault and Giolas interviewed experienced CAD designers to shed light on current industry practices [2005]. Several interviewees believed that sketches (CAD sketched profiles) provide the best method to convey design intent, however there is a tradeoff between complicated sketches and history tree clarity in order to reveal relationships between features.

Zhang and Luo state that CAD illustrates design intent through its history, features, parameters, and constraints [2009]. They state that design intent not only describes an artifact's requirements and constraints, but can also serve an expectant role in the design process. Their research examined methods used to share design intent information between models, but encountered difficulties resulting from an absence of standards and data-exchange procedures.

CAD software manufacturers specify the concept of design intent differently. Siemens' NX [2015], for example, infers that design intent can be extracted by associative parameters, expressions, and constraints so that predictable modification can be achieved. In "history-free mode," design intent guidelines are contingent upon prevailing geometric interactions [Siemens, 2015]. SolidWorks [Dassault Systemes, 2015] defines design intent as "…how your model behaves when dimensions are modified." PTC Creo [2015] provides a definition of design intent where the knowledge of the artifact can be obtained by means of parametric and spatial relationships that define the purpose and fit of the part. These definitions do not represent the conventional understanding of this term. Furthermore, oftentimes problems with the software itself generate a lack of design intent communication, as detailed by Bodein et al, where reusability may be hampered when the relationship between constraints and history is not explicit [2014].

Despite the differences among the various definitions of design intent, it is generally agreed that it is difficult to convey design intent through CAD models. As a result, scholars rely on different methods to communicate this information to others. Some believe that the parametric modeling software can record these data [Rynne and Gaughran, 2008], but while the software can indeed reflect the specific steps taken to create the artifact, it cannot relate why certain commands were used (e.g. when different alternatives exist to create the same feature, why was extruding a profile considered a superior method to revolving it?).

Ault declares that design intent can be acquired by equations used to impose geometrical restrictions based on functional requirements of products [1999b]. Branoff et al. state that dimensioning and geometric relations are devices for establishing design intent within the CAD model [2002]. Bodein and colleagues claim that it is unacceptable that CAD software should provide no capability for the designer to include comments, which are needed to reduce design time [2014]. Dorribo-Camba et al. echoed these thoughts by stating that design intent is often embedded in the modeling approach and in the dependencies between features in the CAD software [Dorribo-Camba and Contero, 2014] and [Camba et al., 2014b]. Their research details methods to use annotations for enabling increased design communication. These annotations are then housed within the CAD model and can be integrated in a Product Lifecycle Management (PLM) system.

The complementary approach to explicitly conveying design intent is trying to automatically detect implicit design intent. Li et al. researched methods to detect design intent primarily by using symmetry [2010a]. They emphasized the identification of design intent by locating prospective geometric abnormalities. Li and colleagues state that geometric constraints and associations between edges, faces, and dependent geometries in CAD models can properly articulate design intent. Their work focused on models bounded by planes, spheres, and cylindrical surfaces, but did not include common curved geometries. Plumed and colleagues researched methods to determine design intent embedded in 2D sketches [2012]. A drawing can be dissected into features, and analysis of such combinations of features can illuminate design intent. The most common features can then be catalogued and identified. Continuing research attempts to examine the feasibility of creating algorithms, which mimic designers' experience and knowledge to extract design intent from sketches.

Even when commonalities exist between various definitions of design intent, oftentimes the manner in which it is assessed (if it is even assessed at all) is inaccurate. To name but one example, design intent that is judged purely by quantity metrics (such as the amount of features or sketches in the design tree) is inherently inaccurate. While quantity metrics are intuitive and easy to calculate, their results may be a poor measure of the model efficiency, as their contribution is nonlinear. When the count is low, the addition of one more is significant, but as the count increases the overall significance of each new item decreases [Owensby and Summers, 2014].

3.1 Representation Structures of Design Intent

As illustrated in Figure 2.1, one of the primary ideas in Design Rationale is that representation allows communication. This concept should be further taken into consideration in order to maximize the resources provided by CAD systems to explicitly communicate the intention of the CAD user during modeling tasks. For instance, proper labeling of modeling operations is clearly a simple way to convey design intent (Figure 2.3), as it eases redesign, analysis, and reuse of CAD models.

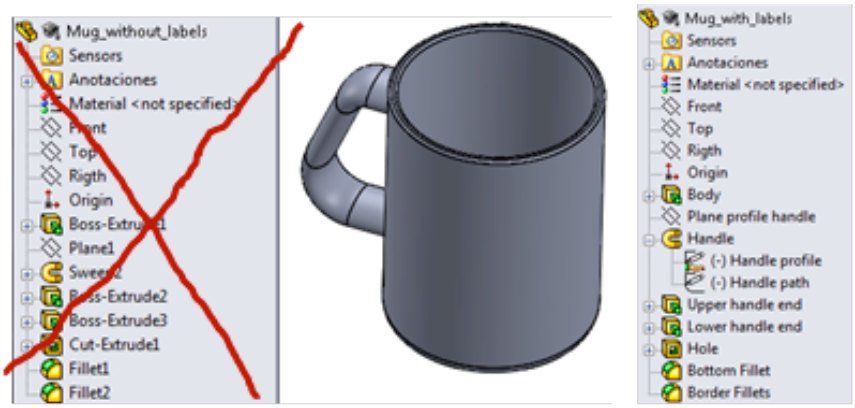


Figure 2.3. Proper labeling of modeling operations as a simple way to convey design intent. Default naming (right tree) vs. proper naming (left tree).

By examining the common structure of commercial MCAD systems, it appears that the design tree is the key representation tool used to manage design intent communication. It is observed that design intent may be embedded in three different model data containers (or structures):

1. Sketches.
2. Reference datums.
3. Modeling operations.

Interactions between all three containers are conveyed through the design tree, which manages and reflects their mutual associations. The history tree not only expresses relationships between modeling operations, but also between datums, datums and operations, and datums and sketches. If it is important to emphasize the sequence of steps, then the term *stages* can be introduced, as sketches necessarily precede modeling operations. But *levels* can also be discussed, since although the container borders are relatively diffuse, each container works primarily at a different semantic level. In general, design intent conveyed by sketches is located at a lower semantic level than design intent conveyed by datums and modeling operations.

In certain situations, all three alternatives are available to express design intent. For example, in Figure 2.4, a simple cylindrical casing is modeled using three different approaches, each of which embeds design intent information in a different container. To guarantee that the cylindrical shape of a casing and its internal hole are concentric, the first approach would link two circles through a concentric constraint defined at sketch level before producing the casing by an extrusion operation. The second approach would result from first extruding the cylindrical shape, then producing the coaxial hole through the cylinder. The two approaches work in a different manner (sketch container for the first approach and modeling operations for the second). But in both cases, the approaches assume that cylindrical shapes are to be obtained from extruding circles. This ambit reveals a line of thought where the mind's eye first works in 2D (circle) and then extends the result up to 3D (cylinder). However, a cylindrical shape may also be obtained through a revolution operation applied to a generatrix. In our example, the casing shape may be obtained by revolving a slender rectangle around an external axis of revolution. This option (which changes the strategy at the third level) is less intuitive, as we only see a rectangle, but it explicitly defines the axis that is (a) single and (b) shared by both generatrices. As a result, the cylindrical shape and the cylindrical hole are necessarily coaxial.

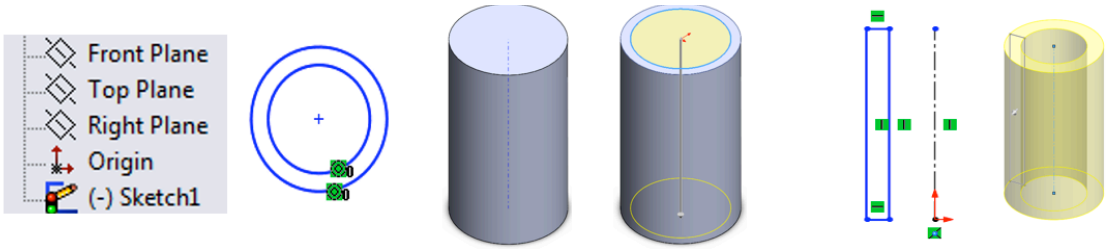


Figure 2.4. Three approaches to model a casing: (a) Sketches (ex: concentric link between circles at revolve sketch level plus extrusion.), (b) Relationships between Modeling (ex: construction of cylinder followed by construction of coaxial hole.), and (c) Modeling Operations Itself (ex: Profile of casing wall and profile to create casing (revolve)).

Selecting the best of the three data containers to introduce design intent in CAD models requires agreement between contradictory criteria; because the three containers described above have different visibility. They also differ in ease of being defined and modified. For example, adding constraints to a sketch is fast and reliable, but hides the design intent within the sketch, which is not directly visible when inspecting the design tree. This issue should be considered when it is possible to choose between multiple representation stages. Choosing those alternatives that are directly visible on the design tree facilitates future analysis and understanding of the CAD model, but may also be inefficient in certain situations.

One example of hidden references is shown in Figure 2.5, where the reference axis (surrounded by a lasso) is embedded in the base profile, where unnoticed changes in the profile may inadvertently change the scaffold.

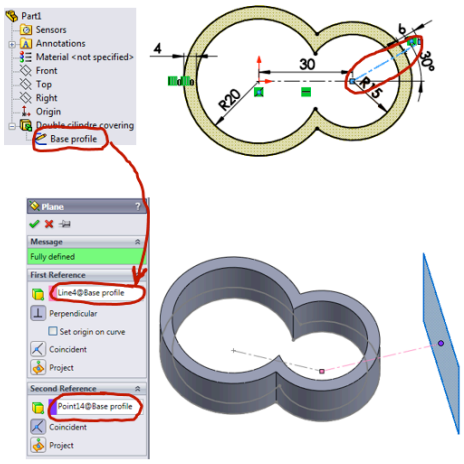


Figure 2.5. Example of reference datum supported by one line embedded in a profile.

Enhancing the visibility of design intent requires expert users capable of working with sophisticated modeling operations, which greatly affects model portability. For example, adding constraints at the 3D level is somewhat more limited than the 2D alternative, since available constraints are usually reduced to bilateral symmetry and rectangular and polar replication patterns. On the other hand, these 3D transformations help improve the visibility of design intent.

The 3D approach is also more sensitive to round-off errors. For instance, adding a symmetric element to a body by a bilateral symmetry operation (as shown in the second lug in Figure 2.6) may unintentionally result in a multi-body model, simply because round-offs in the size or the location of the symmetric element (more likely to occur in 3D calculations than in 2D) may produce small cracks that prevent the new element from being merged to the main body.

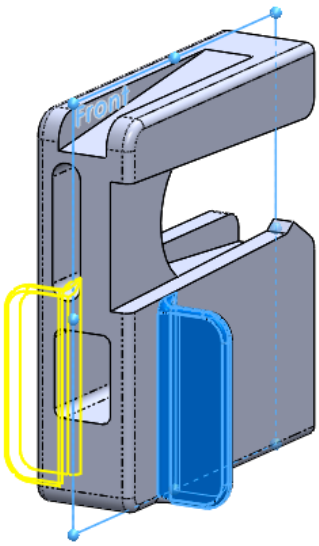


Figure 2.6. Addition of a second lug using symmetry.

The embodiment of design intent is associated with, "readability, alterability, and usability of CAD models [Mandorli et al., 2016]." Mandorli and colleagues state that engineers are able to deduce design intent by the artifact's shape, dimensions, and tolerances and are interested in translating design intent representation from historical 2D drawings into 3D objects [2016].

Explicit, or at least easily accessible, design intent communication is a first step toward expressing design rationale in a proper manner and supporting the creation of CAD models that are ready for redesign, analysis, and reuse. It is essential for CAD users to realize their modeling decisions must express, as clearly as possible, what the expected behavior of the model should be. The most appropriate representation level of design intent (sketches, relationships between modeling operations, and the modeling operations themselves) must be selected; so all information is communicated effectively. Furthermore, explicit communication of design intent must be a core ingredient of CAD instruction.

3.2 Capture and Transfer of Design Intent

As illustrated previously in Figure 2.1, one of the main positions in Design Rationale is that it responds to the idea that representation allows for communication. Hence, modeling strategies should be aimed at maximizing re-design, analysis and re-use of CAD models. Authors with comparable definitions of design intent rely on different methods to communicate this information to others.

There has been much research and discussion, from the early development of parametric modeling, about how best to capture design intent. In general, researchers realize that the extraction and sharing of design intent is crucial [Will, 1991 and Shih and Anderson, 1997], but previous attempts have failed because of incompatibilities in software and inefficient storage methods [Anderson and Ansaldi, 1998].

An early examination by Will [1991] focused on industry use of simulation and modeling and concluded that access to past designs and their design intent was crucial to industrial product development. He also believes that this information needs to be recorded and placed in libraries so that engineering changes can be made using this data. Shih and Anderson, while investigating product model data sharing, state that if this data cannot be captured and retained, a barrier exists preventing the exchange of product information between designs [1997]. In a continuing study, Anderson and Ansaldi state that data exchange between CAD software is problematic, because they use different constraints and algorithms in the solvers [1998]. In addition, CAD vendors would resist standardization of the solvers in order to protect their proprietary systems [Anderson and Ansaldi, 1998].

To alleviate data exchange problems, various software solutions have been suggested. Anthony et al. describe an approach to use Conceptual Understanding and Prototyping environment (CUP) to capture design intent [2001]. CUP documents design information about assemblies and stores it using concepts characterized by XML [Anthony et al., 2001]. Similarly, Choi et al. suggest using macro files [2002]. In their study, commands are grouped into categories and ACIS 4.0 (a geometric modeling kernel) is used to generate an internal geometric model in order to overcome unresolved command mapping between software [Choi et al., 2002]. Fu et al. state that development of software-independent tools is necessary to manage CAD data and the conversion of design intent into process-applicable information [2002].

Further research points toward CAD data exchange standards in order to facilitate design intent transfer. However, the main issue related to this approach is that the more widely used neutral formats such as IGES or application protocols 203 and 214 of ISO 10303 (STEP) do not support the transfer of sketches or modeling operations (features) and their relationships, as recognized by Pratt et al. [2005]. In addition, design intent information is lost during file transfer between systems because STEP does

not allow representation of it [Pratt et al., 2005]. This situation has started to change since the 2014 publication of the first edition of Application Protocol 242 (AP242), "Managed model based 3D engineering." Valid implementations of this protocol are beginning to appear, which are expected to alter the situation drastically. This new application protocol supports all the elements mentioned previously, however, there is no current commercial CAD application that supports the units of functionality related to parametric history-based feature-based modeling. According to Chang, a plan for incorporating design intent, features, and profiles with corresponding constraints and dimensions is highly beneficial, especially for new CAD users [2014]. Pressure from end-users and industry should finally force CAD developers to implement the whole AP242 capabilities fully.

Another important aspect related to the transfer of design intent is that oftentimes the lack of tools for the visualization and analysis of relationships between features in CAD models presents an important barrier for understanding design intent. As detailed by Bodein et al, reusability of CAD models is hampered in that the relationship between constraints and history is not explicit [2014]. They also declare that the history tree does not adequately express the relationship between features, especially if certain features are created in a nonlinear process [Bodein et al., 2014].

3.3 Design Intent Instruction

Effective learning of MCAD tools remains a major challenge in both academic and industrial settings, with most educational practices focusing on declarative and specific procedural command knowledge, as classified by Chester [2008]. Declarative command knowledge is related to generic commands or algorithms that are typically available within MCAD systems, such as geometric transformations (i.e. patterns and mirroring operations), or basic solid modeling operations (i.e. extrusions or revolutions). The practical application of this realization to a particular CAD system requires specific procedural command knowledge, which is system and version specific. Usually, procedural command knowledge occupies the bulk of the teaching/learning activities in a MCAD course. However, this approach does not provide sufficient strategic knowledge, which is associated to the election of the proper modeling strategies. The concept of quality in the context of MCAD systems is also omitted. In this context, the understanding of the design intent concept by the trainee is critical to develop strategic knowledge of the MCAD tool.

Research has been performed on methods to increase the amount of design intent available for communication, with much of this effort targeted at beginning CAD learners. Condoor states that historically there was one correct depiction of an artifact [1999], but with CAD, that artifact may be created using several different approaches, with some techniques being superior in that they more successfully reflect design intent. He determined that there is a substantial connection between the methodology used to create models and the inherent design intent and proposed a procedure to instruct CAD learners to better reflect design intent by subdividing assemblies into parts, and parts into specific entities; identification of symmetry; proper datum plane orientation; design sequence; and hypothetical changes [Condoor, 1999].

Hartman, in a two-part study attempting to determine how experienced CAD designers achieved their current level of expertise, states that new CAD learners need curriculum that provides instances where models are created, altered, and model geometry can be manipulated so that they can be adequately prepared for real-life design complexity. Curricular exercises need to be created so that the correctness and acceptability of an artifact can be related to the model's response to future design changes, both expected and unexpected [Hartman, 2004 and 2005].

Johnson and Diwakaran claim that while rapid model creation is valued, creating designs quickly adversely affects design intent and model perception [2009]. They assert that the quality of a model should correlate with the amount of time needed for revision, which attempts in some way to quantify design intent and its communication between users. In a continuation of their research, Diwakaran and Johnson conclude that CAD models must be easy to change so that design alterations in the product

development cycle are accomplished quickly [2012]. It was determined that using simpler features increases the time required to model the original artifact, but increases the reuse of the model in future incarnations. Additionally, simple features, along with the use of reference datum and correct feature sequence increase model understanding when undergoing alteration by secondary users. Feature alteration and reuse is positively correlated with model perception.

Similarly, feedback and evaluation have also received attention. Leahy suggests that well-timed feedback of student performance is needed so that students can incorporate best practices for design intent [2012]. He suggests that feedback be non-graded in order to encourage students to strive for deeper knowledge instead of being motivated exclusively by higher marks [Leahy, 2012]. Ramos-Barbero and Garcia-Garcia echo a similar pedagogical philosophy, stating that student errors should not be considered failure, but a natural part of the learning process in that it highlights the importance of design methods and standards when using CAD [2009].

Proper model assessment for design intent communication is an arduous task, especially for large class environments. Branoff and Wiebe claim that evaluation of student work in order to assess for proper levels of design intent requires accessing student models, which is a time consuming task and oftentimes leads to examination of the hard copy [2009]. To alleviate this problem, tools to convey feedback and/or evaluation have been suggested. Baxter and Guerci developed a method to automate the assessment of CAD models and give students instantaneous evaluation [2003]. Macros were written and used in conjunction with the Application Programming Interface (API) associated with the CAD software to facilitate the uploading of models to a database. The API facilitates a grading program that compares data from the student files to a master model provided by the instructor. But this approach hardly evaluates the amount of design intent conveyed by the CAD model.

Kirstukas developed a computer program that evaluates the geometry and alterability of student solid models [2016]. This program compares student models against an instructor-provided one, deducting points for unused CAD sketched profiles, non-united bodies, and banned constraints while calculating various mass properties [Kirstukas, 2016]. But as with earlier attempts at automating model assessment, design intent and how elegantly the artifact was constructed are not extracted.

Irwin examined what he named scaffolding techniques (mentoring students toward finding solutions while adjusting the amount of support provided based on the level of student performance) in a senior-level design course to optimize CAD model usability [2013]. Constant values in the CAD model were replaced with expressions, which drive design intent, allowing for increased flexibility of design exploration. This approach is extremely valuable because it emphasizes the importance of requiring models to be controlled by parameters (not just linear dimensions) to drive design intent and allow for increased flexibility of design exploration [Irwin, 2013].

Ramos-Barbero and colleagues determined that students with stronger spatial vision applied design intent strategies better [Ramos-Barbero et al., 2016]. They also state that CAD model alteration should be integrated early in CAD instruction so learners will understand appropriate modeling schemes, while additionally calling for improved design intent rules for assemblies [Ramos Barbero et al., 2016].

Camba and colleagues examined different methods to create reusable 3D models [2014a]. They found that “Horizontal” modeling provides for easy alteration because features are independent elements, but that this method is not the most intuitive strategy and describe it as producing the most flexible models, but reduces the functionalities that make the model parametric [Camba et al., 2014a]. They further state that “Explicit References” modeling provides for simple models but are difficult to model and “Resilient” modeling is effective, in spite of reference nodes needed to reduce dependencies [Camba et al., 2014a]. In a more detailed examination, these different modeling methodologies were studied to examine CAD reusability [Camba et al., 2016]. “Horizontal” modeling minimizes CAD repair by removing parent/child dependencies between features, “ Explicit References” modeling minimizes the number of constraints linked to existing geometries by managing functional references, and “Resilient” modeling manages the sequence of the design tree by organizing the features by

purpose and priority [Camba et al., 2016]. The researchers found that "Resilient" modeling was most effective, although the level of CAD expertise of the study population (students) could have affected the results [Camba et al., 2016].

Goodrich defined rubrics as assessment tools that specify important curricular concepts as well as gradations between quality levels [1996]. Devine and Laingen implemented an assessment scheme that utilized grading rubrics, feedback, and model manipulation to verify design intent [2013]. Company et al. state that students need explicit procedures and metrics to assist them in evaluating their performance and describe an expand-contract approach to convey quality-oriented strategies to CAD trainees by embedding quality criteria into rubrics so as to force CAD trainees to understand them early in their instruction [Company et al., 2013 and 2015]. In these rubrics, students' work (a CAD model or assembly) is broken down into its components, that are checked against quality dimensions (conveyed as competences), and later measured through evidences or "assertions [Company et al., 2015]." Design intent is addressed in one of the dimensions covered by the rubrics proposed by these authors.

The catalogue of methods utilized to increase the amount of design intent is incomplete and will remain so. Nevertheless, detecting as much common intent as possible is still feasible and useful. Feedback and assessment of the design intent conveyed by a CAD model is also an open problem. Based on current research, it is believed that rubrics and assertion maps are a promising approach, especially when specific quality dimensions are related to the proper communication of design intent. In some cases, the process of validating assertions can be automated, which provides new opportunities in the field of intelligent tutoring systems applied to CAD learning.

4. Conclusions

Although design rationale is a well-established field of study, in the context of MCAD systems, design intent remains a complex concept, with different visions and approaches available in the scientific literature. Design intent is commonly, but not always, understood to describe a model's anticipated behavior once it undergoes alteration. There is a consensus that modeling tools and strategies greatly influence design intent communication. There is also agreement in the convenience of expressing design intent through proper modeling strategies, especially when beginners are learning to model.

Strategies and approaches aimed at improving expression of design intent into CAD models to enhance their quality, together with metrics aimed at evaluating its efficiency, are now receiving some attention. It is becoming evident that guidance aimed at specific design intent instruction is required, since it has been argued that enhancing design intent conveyed through CAD models may be performed at three different stages (sketches, reference datums, and modeling operations themselves), which have different advantages and disadvantages that must be balanced to get an agreement which allows for selecting the best modeling strategy.

Research shows that rubrics can be a useful tool to facilitate standardized design intent communication. Rubrics are important not only for assessment, but also for communication of expectations. Of current interest is how to define qualities of design intent (and model quality) in such a manner that lends itself to easy assessment. More precise definitions of these terms are vital to any productive research being accomplished. Further development of these concepts to construct assessment rubrics with the goal of standardizing such definitions and assessment strategies are envisioned. These rubrics must be adaptive towards the individual and his state of knowledge and other preferences (rubrics change in a system-driven base). They must also be adaptable, as their personalization must be controlled and steered by the user (i.e., user-driven). CAD model quality should not be a correlative goal only to be attempted after basic skills are cultivated, but a major goal from the inauguration of instruction.

Chapter 3
The Process for Assertions Maps and Rubrics Definition

1. Introduction

This chapter describes an approach to convey quality-oriented strategies, organized around quality dimensions, to CAD trainees by embedding quality criteria into rubrics so as to incorporate these strategies early in the instruction. Subsequently, an approach is introduced based on progressive refinement, which results in an assertions map that indicates quality dimensions vs. sequence of tasks for CAD models. This assertions map illustrates how the expand-contract strategy adapts the rubrics to CAD trainee progress and assists them in comprehending the different rubric dimensions. Also highlighted are specific insights gained on the suitability of using separate rubrics for different tasks, the need of accurately timing the expand-contract process, and on the convenience of supporting rubrics with appropriate instruction, focusing on the conveyance of good practices and evaluation tools through rubrics. In summation, it is proposed that quality rubrics must be both adaptable (user-driven) and adaptive (instructor-driven).

Currently, New Product Development (NPD) processes are entirely based on digital product representation, from 3D definition to digital manufacturing. Their implementation of collaborative/concurrent engineering is exclusively centered on Product Lifecycle Management (PLM) systems that provide support for the exploitation of 3D CAD models along the NPD process. Standards relating to digital product definition data practices such as ASME Y14.41-2012 and ISO 16792:2015 are reinforcing the central role of 3D CAD models in this context.

The growing importance of 3D CAD models as the central core of the NPD process has been accompanied alongside the corresponding development of the CAD data quality field. Standards, such as the "Strategic Automotive product data Standards Industry Group" (SASIG) Product Data Quality (PDQ) Guideline V2.1 (ISO/PAS 26183:2006) and VDA 4955/4.1, provide basic quality criteria for "product data" that is defined as any and all product data required from product conception to manufacturing. Product data includes not only computer aided design (CAD) data, but also data generated from computer-aided manufacturing (CAM), computer-aided engineering (CAE), and product data management (PDM), among others, and CAD education should address all phases of the life cycle [Dankwort et al., 2004]. However, current versions of these standards are specifically oriented to provide mathematical and topological quality criteria for CAD data.

Poor data quality management represents a serious hurdle to obtain all promised benefits of history-based parametric feature-based mechanical CAD systems (MCAD), provoking avoidance, mitigation, and delay costs [Brunnermeier and Martin, 2002]. These concerns are especially critical regardinAD model reuse issues. Partial redesign of an existing product is a typical approach in the NPD process [Ullman, 2003]. Here, high quality CAD models play an important role, as reported by the Aberdeen Group in 2007 [Jackson and Buxton, 2007]. Companies reported a 30% reduction in design times for new products that were closely associated with preexisting products, creating an 80% reduction in design times for products heavily dependent on reuse. However, reusing CAD models is not without problems, as the same report [Jackson and Buxton, 2007] indicates the following obstacles to CAD model reuse:

- Model modification requires expert CAD knowledge.
- Models are inflexible and fail after changes.
- Only original designer can change models successfully.

These problems are not new and are continuing. In 1998, Anderl and Mendgen [1998], in researching the creation of real life complex CAD models said, "If it is difficult to create a model then it is even more difficult to reuse it for variation of modification purpose." Horwood and Kulkarni [2005] indicate that defective data can be attributed to factors such as improper modeling practices, lack of communication of methods, nonconformance to generic methodologies, neglecting the enforcement of quality standards, time pressures to complete the work, too many data translations, and lack of training and oversight.

In another study, Bodein and colleagues [2013] identified as the main aspect axes for an efficient CAD strategy:

- To reduce design time in all design phases (conceptual, preliminary, or detailed),
- To reuse existing CAD models and geometry,
- To accelerate the automation of routine design tasks based on knowledge-based engineering (KBE) applications,
- To enhance collaboration between designers, and
- To improve the general quality of CAD models.

Previous research suggests a proactive approach to embed quality concepts from the initial stages in CAD training in an explicit way. This concept is aligned with the idea by Bhavnani et al. [1999] that strategic knowledge holds the key to efficient CAD usage and that this knowledge must be explicitly taught. Strategic knowledge can be defined as that knowledge related to identify alternate working procedures and how to choose between them in order to make the most efficient use of the CAD tool.

Rubrics have been successfully tested as a valid tool for disclosing and conveying those quality criteria that can be efficiently transmitted by way of good practices in a basic 3D CAD course for fresh designers/engineers [Company et al., 2013]. Additionally, it has been reported [Company et al., 2014] that only specific rubrics are useful and varying levels of detail are required at different phases in the training period. Hence, in this chapter, a study of the organization of quality criteria in rubrics is performed.

An approach to develop quality-oriented rubrics is described, based on a panel of experts that progressively refine quality criteria that can be disclosed and conveyed through rubrics, while determining when and how they can be transmitted to the CAD trainees. Instead of constructing a single rubric, this approach intends to produce what are called Assertions Maps. These maps are visual representations which indicate which assertions are necessary to include in the different rubrics used to convey quality in CAD documents throughout the training period. The assertions maps display the

expand-contract process, where the assertions are significantly detailed the first time they are introduced, and are later recursively abstracted to force CAD trainees to comprehend general quality criteria, as suggested in [Company et al., 2014]. To validate this approach, three pilot experiments were conducted.

2. The Approach to Develop Rubrics

Rubrics are a common authentic assessments tool to describe student achievement [Goodrich, 1996 and Goodrich, 2000]. A rubric is a scoring tool that lists the criteria for a piece of work and articulates gradations of quality for each criterion. Rubrics can increase student performance by making explicit the instructors' expectations and by showing students how to meet these expectations. Rubrics are also useful to assist students to become more thoughtful judges of the quality of their own and others' work. True assessment emphasizes the application and use of knowledge to solve complex tasks that involve contextualized problems. Rubrics are an invaluable tool in order to help students understand the criteria for judgment from the inauguration of their instruction [Montgomery, 2002]. All these arguments indicate choosing analytical rubrics (rubrics that decompose a work into its components that are judged/scored separately first and then later combined to produce the evaluation) to both communicate and evaluate CAD quality criteria. Using rubrics to convey quality is surprisingly straightforward: quality dimensions are conveyed as rubric competences, and competences are measured through evidences, which are expressed as "assertions".

Not all quality concepts can be introduced simultaneously, however. So, the proposed approach uses assertions maps that are visual representations of sets of rubrics. The method to determine these assertions maps is simple but efficient: a panel of experts recirculate the proposal, which is recursively refined. In other words, instead of directly testing the rubrics with students, a team of teachers/researchers iteratively polished them. The approach consists of six stages.

First stage:

1. Quality concepts are organized into dimensions.
2. The team agrees on a course syllabus and the quality concepts that are embedded in it.
3. The syllabus is organized around tasks structured in series.
4. Assertions maps are obtained to convey the expand-contract process.

Second stage:

5. One team member develops a particular task and its associated rubric.
6. Another team member solves the task and satisfies the rubric.
7. Both members discuss the weak and strong points of the rubric as a means to convey quality criteria and produce an improved version.

Third stage:

8. One member of the team develops tests to quantitatively validate those assertions that are difficult to validate directly.
9. Another team member uses the tests to satisfy the rubric.
10. Both members discuss the weak and strong points of the tests and improve them.

Fourth stage:

11. While repeating stages two and three in order to accomplish all tasks and their associated rubrics, the rubrics are compiled in the assertions map and their suitability for the expand-contract criterion is checked.

Fifth stage:

12. Post-graduate students are asked to comment on the clarity and usability of rubrics.
13. Collected comments are used by the experts to produce an improved version.

Sixth stage:

14. The rubrics are finally tested with undergraduate students.

A simple example illustrates this process: creating profiles is a mandatory task for all sweep-based 3D CAD applications. Robustness is defined as a profile's quality that is clearly related to its constraints. Therefore, drawing a fully constrained profile is a pertinent, quality-oriented task. The rubric used to assess the task should include an declaration such as, "The profile is fully constrained." This statement should first appear as a separate entry at the beginning of the assertions map and should be grouped with similar asseverations (such as, "The profile does not contain duplicated lines" and "The profile does not contain segmented lines"). The aggregated pronouncement should be, "Profile is robust". Finally, all assertions should be implicitly embedded in the general dimension of model consistency.

This approach is similar to the "learning to see" methodology proposed by Bhavnani et al. [1999], but instead of simply focusing on developing the teaching of strategies, the proposed approach guarantees that CAD trainees are forced to acknowledge the importance of good practices, as they are explicitly required to verify whether they have accomplished them. Furthermore, the concept of "learning to do" is extended as simple tests to check the quality concept under evaluation are developed. As an example, in order to detect whether the profile is suitably constrained, students are required to edit one of its dimensions and check whether the result exactly matches a given template.

Sections 3 and 4 of this chapter describe the background and development of the first stage of the proposed process. Examples that illustrate the fifth and sixth stages are detailed in Section 6, where the experiments describe the different lessons learned and illustrate the suggested process for replicating the approach for further developing rubrics adapted to other scenarios.

The second, third, and fourth stages are not further clarified, as they are primarily based on applications of team member expertise. They must be solved using what Rossignac defined as education-driven research (EDR), which develop "specific tools and solutions in a specialized domain, so as to make them easy to understand and internalize" [2004].

CAD education was identified by Piegl as one of the ten challenges in computer-aided design [2005]. He suggested that the internet will facilitate the re-use of CAD components, especially with the interconnection of computers and access to databases. It is foreseen that models must enforce re-usability (and, in general, quality) before becoming publicly available. He also predicted the advent of on-line CAD education. It is anticipated that self-evaluation tools will be required, and suggest using rubrics toward that end.

According to the literature, at least three main topics must be considered in CAD education: (1) industrial view, (2) teaching of fundamental tools and knowledge, and (3) teaching of more advanced and specific topics [Sapidis and Kim, 2004]. This chapter emphasizes the industrial view, which includes balancing the mathematical foundations of CAD, its relation to computer science, design

methodologies, and system evaluation [Ye et al., 2004]. This work addresses one particular design methodology: feature-based modeling.

The continuing evolution in the training and educational needs of users of such CAD-systems was already highlighted by Field, who distinguished three groups of users: the majority, the expert, and the super-user [2004]. This chapter spotlights the approach of using rubrics to shift users from the majority group to experts, in the sense that they become competent to edit and repair CAD models and transfer information to and from different secondary views (Finite Element meshes, etc.). As expressed by Dankwort et al., CAD-education is not restricted to only teaching solid or surface modeling, but includes learning about the complete development process under the aspect of Computer Aided Product Creation [2004].

Hamade and Artail distinguish four types of students (Activist, Pragmatic, Theorist, and Reflector), and describe tests to correctly place trainees in groups, arguing that this methodology increases the efficiency and cost savings of the learning process [2008 and 2010].

Amadori et al. [2012] state that if users are taught to use efficient strategies in the context of tasks, they will be able to recognize opportunities to use them in new scenarios. Mandorli and Otto recognize the importance of forcing CAD trainees to comprehend quality criteria early in their instruction in order to prevent the Einstellung effect (the consequence of past experience negatively effecting the ability to resolve fresh problems) [2013].

To sum up, between the second and fourth stages of the approach, experts should remember that CAD education should take into account all aspects of Computer-Aided product creation, distinguish the four types of students, and pivot around diverse tasks.

3. CAD Quality and its Dimensions: Related Work and Model Proposal

After a detailed analysis of standards on product data quality and working procedures in the industry related to CAD data exchange agreements, three “levels of quality” can be identified [Contero et al., 2002]. The morphological quality level is related to the geometrical and topological correctness of the CAD model, and currently it is the focus of PDQ standards such as SASIG PDQ Guideline V2.1 and VDA 4955. The syntactic quality level is linked to the proper use of modeling conventions such as naming rules for features, datum, part, assembly, drawings and layouts; layer structure and function and part/assembly parameters and attributes. The syntactic level is especially focused on the organizational aspects of the CAD model. Finally there is a third level associated to the semantic/pragmatic quality that takes into account the CAD model capability for reuse and modification. CAD users have an abundant variety of modeling procedures for shaping their designs. However, experience shows that certain procedures provide better solutions than others.

Knowledge linked to the semantic/pragmatic quality level corresponds with the strategic knowledge, defined by Bhavnani et al. [1999] as knowledge for identifying alternate working procedures and selecting the one that most efficiently uses the CAD tool. Also, it can be identified with the procedural knowledge, as opposed to declarative knowledge by the following researchers [Rynne and Gaughran, 2008], [Johnson and Diwakaran, 2009], [Diwakaran and Johnson, 2012], and [Mandorli and Otto, 2013].

Making this knowledge well documented and easily accessible is extremely important. Many large companies have developed their own private "modeling guidelines", where the "best practices" are recorded for improving CAD model quality. Some authors have proposed modeling methodologies to improve overall quality in CAD models. Rynne and Gaughran [2008] termed it “cognitive modeling”. The most recent contributions are the “resilient modeling strategy” by Gebhard [2013], which advocates replacing best-practices manuals with checklists and the “explicit reference modeling methodology” that proposes a modeling methodology for complex parts based on explicit references

[Bodein et al., 2014]. Other modelling methodologies, such as "horizontal modeling" are protected by patents [Landers and Khurana, 2004].

A data quality dimension [Wang and Strong, 1996] is defined as a set of data quality attributes that represent a single aspect or construct of data quality. This chapter intends to establish the dimensions of the CAD quality space that can be introduced while training novice CAD users. This scope is limited, as this work does not intend to encompass all aspects of quality in CAD models. The selected dimensions are inspired by the properties of representation schemes proposed by Requicha [1980], where formal properties are: domain, validity, completeness, and uniqueness. Informal properties are: conciseness, ease of creation, and efficacy in the context of applications.

The domain of representable entities is a prior limit of each CAD application, which excludes certain categories of shapes, but does not influence the quality of those designs that fall within the range of the application. Similarly, the ease of model creation relies on the friendliness of the user interface, which cannot be readily improved by an experienced user. Both properties should be considered only if users were allowed to select between different CAD applications in order to solve each particular task.

Validity is a simple property, but an important concept while training novices. Solid models are electronic documents that must be correctly saved for later use. A common mistake when beginning to use CAD software forgetting that an opened file must not be manipulated by the Operating System. Attempting to save files while they are still in exclusive use by an application usually produces invalid documents. Furthermore, mistakes made while modeling result in model trees with errors that produce non-usable files.

Models are complete if they include all the product aspects that are relevant for design purposes (that is, replicate the shape and size of the object). Completeness appears to be an obvious requirement and also implies quality, as incorrect modelling flow frequently produces a non-complete model.

Uniqueness of the internal representation is a simple way for assessing the equivalence of objects. However, modern parametric feature-based systems prioritize ease of use, allowing different construction routes to achieve the desired geometry. This ease of use compromises quality, since as noted previously when the idea of "semantic/pragmatic" level of quality was introduced, different paths to produce the same geometry react differently to CAD model changes or modification. In this context, consistency expresses in a more exact way than uniqueness, the quality dimension linked to the goal of allowing safe exploration of alternative solutions while maintaining valid geometries.

This research is not interested in the efficiency of the modeling process itself (i.e. how quick or simple the modeling task is for the designer). Instead, the quality of the high level information conveyed through the resulting model (i.e. how well does it represent the object, how easy it is to alter, or how much design intent it conveys) is of interest.

Hence, this research seeks to determine whether the actions of the CAD user (what tasks the user performs during the modeling process), result in documents that convey design intent. Extensive discussion of the meaning and scope of design intent is included in Chapter 2, but for the sake of brevity, two similar definitions that explain what are understood by design intent include: a) CAD model's capability of being modified and yet still being able to perform the same functions, and b) expected behaviour of a CAD model when it is modified. Therefore, measuring whether the modeling process is effective in conveying the right information about a model's function must be attempted first. Additionally, of interest is whether measuring the modeling process conveys the information in a correct manner: if the modeling process allows conveying more design intent than other modeling processes (i.e. is efficacious in conveying design intent), and if the modeling process requires less effort than others to convey the same design intent (is efficient).

So, a classification system is developed with the following dimensions of quality in CAD models:

1. Models are valid if they can be opened by suitable applications and do not contain errors or warnings.
2. Models are complete if they include all the product aspects that are relevant for design purposes.
3. Consistent models should not crash as a result of editing tasks and design exploration should be easy.
4. Concise models do not include irrelevant or repetitive information or procedures.
5. Effective CAD models convey design intent.

This classification can be confirmed and improved by reviewing recent literature. Bhavnani et al. [1999] place a premium on efficiency. Rynne and Gaughran [2008] defined a set of attributes which cover the dimensions of efficiency, robustness and proper design intent.

According to Diwakaran and Johnson, the ability to create and easily alter designs is one of the key proposed benefits of CAD programs, but models cannot be easily altered by other designers and engineers if these models "are difficult to understand or do not capture design intent well" [2012]. In fact, McKenney reported that many engineering analysts are spending half their time repairing poorly constructed CAD models before analysis can begin [1998]. Hence, the four main dimensions of a quality CAD model should be: ease of creation (speed), ease of alteration, ease of understanding, and effective capture of design intent.

However, as stated before, it must be understood that speed is not a quality goal. In fact, Hamade et al. [2005] concluded that if production time is an overriding criterion, then using small numbers of complex, more time-efficient features is the best choice. However, speed contradicts ease of alteration. Diwakaran and Johnson concluded that relative feature complexity (greater average number of segments per feature) was found to decrease original modeling time, but also found to decrease design reuse (through lower feature retention without change) [2012].

It is noted that those properties encompass capabilities belonging to different dimensions. For instance, strategies to increase feature reuse during alteration are: simpler features, the use of reference geometry, and the correct feature sequence improve the perception of the model during alteration.

While ease of alteration is similar, it is not equal to consistency. In the latter, the emphasis is on the number of changes (the more changes that are allowed the better), while in the former, the emphasis is on the ease of common changes (simpler changes are preferred).

Those dimensions are not independent of each other. For instance, reducing to a minimum the amount of information (i.e. conciseness) is occasionally contradictory with easing the understanding of such information, as understanding usually improves with moderate redundancy.

Such contradictory requirements become even more apparent if it is considered that, according to Amadori et al. [2012], models should be flexible and robust. The wider the range of product configurations, arrangements, and sizes the model can cover, the more flexible it is. The fewer the amount of errors or instability issues that changes to the geometrical model may provoke, the more robust the model. In other words, a CAD model is robust if it can be modified without failure. Also, they established that complexity is also related to flexibility and robustness, since robustness and flexibility values tend to worsen when the complexity of models increases.

Assuming that CAD models are documents shared by different stakeholders during the design process, communication is also important. In order to facilitate communication, the document must follow conventions and must be clear and comprehensible (aimed at being understood at first glance) [Company et al., 2013].

The Resilient Modeling Strategy [Gebhard, 2013] considers that models must be editable-robust, obvious, and reusable. This strategy hypothesizes that parent-child relationships are critical in order to allow robust editing, but are sensitive to the feature sequence. He argues that in order to get robust models, modeling operations should be sequenced exactly in the following way:

1. Reference operations (datums).
2. Construction operations (skeletons, or scaffolds).
3. Core operations (additive features)
4. Detail operations (subtractive features)
5. Modify operations (replication operations)
6. Quarantine operations (cosmetic operations).

The Resilient Modeling Strategy™ [2017] has recently been updated to version 4 and reorganized into the following groups:

1. Skeleton Group (reference geometry).
2. Core Group (additiive features).
3. Surface Group (curves and profiles).
4. Detail Group (subtractive features).
5. Holes Group (cylindrical cavities)
6. Modify Group (replication operations).
7. Quarantine Group (edges).

The author agrees with this sequence, but suggests training CAD users to work with true design or manufacturing features instead of simple form features. Hence, it is emphasized that subordinating the priority additive-before-subtractive to the more important priority of design and manufacturing features before simple form features.

Finally, design intent should be understood as the way to describe a model's anticipated behavior once it undergoes alteration [Otey et al., 2014]. While many authors have comparable definitions of design intent, they each rely on different methods in order to communicate this information to others. It is not believed that parametric modeling software can accurately record this data. Instead, methods need to be developed so that this information can be documented and design justifications understood. Still, some simple actions can be performed with current CAD applications. To develop obvious models, this approach allows for communicating design intent by renaming features and reordering the model tree so that it reads like a recipe for the model.

To sum up, the dimensions of quality, including some sub-dimensions, should be reformulated as follows:

1. Valid
    - 1.1 Retrievable (can be found and opened)
    - 1.2 Usable (is error-free and compatible with the application)

2. Complete

2.1 Replicates shape of the object

2.2 Replicates size of the object

3. Consistent

3.1 Robust (changes do not produce unexpected failures)

3.2 Flexible (allows many changes)

4. Concise

4.1 Non-repetitive (does not contain repetitive information or operations)

4.2 Non-fragmented (does not contain fragmented information or operations)

4.3 High semantic (uses high-level modeling operations when available)

5. Simple

5.1 Clear

5.2 Easy to understand (an observer can easily explain the model)

5.3 Maximizes compatibility with other CAD formats ("saving as" produces files retrievable and usable by other applications)

5.4 Follows conventions (interpretation is non-ambiguous)

6. Captures design intent

6.1 Effective (conveys design intent)

6.2 Efficacious (conveys more design intent than other modeling processes)

6.3 Efficient: it is preferable to model this way in order to convey design intent (the same design intent could not be conveyed in a simpler manner).

4. Quality Oriented Training

It has been stated that rubrics must adapt to the task, and tasks must be arranged in the correct sequence [Company et al., 2013]. As a first step to cope with different tasks, a subdivision between modeling, assembling, and drawing extraction was considered and this subdivision is maintained. Also, the tutorials have been organized into series, where a series is a set of exercises aimed at training one specific skill. Exercises belonging to every series are internally ordered by increasing level of difficulty, from simple repetitions of the basic skill being trained, to using this skill in novel scenarios. The external sequence of the different series is also important because the needs and capabilities of CAD trainees evolve along the training period. To adapt to the evolution of CAD trainees, an expand-contract strategy was proposed. Hence, a sequence of tasks and sub-tasks was developed. Quality concepts are organized around main dimensions, which in turn, are evolved into sets of criteria expressed by way of assertions. Both sequence of tasks and dimensions of quality criteria constitute the two axes of the assertions maps.

The 30-hour course is aimed at teaching 3D CAD modeling fundamentals to mechanical and industrial engineers. CAD trainees had been exposed to a previous course named "Graphic Expression", where they learned the fundamentals of descriptive geometry and standard representation of engineering drawings. They were also instructed in the use of those fundamentals to produce both hand drawn

sketches and 2D CAD engineering drawings. The CAD software used in the experiment was SolidWorks®.

After analyzing alternatives to configure the course syllabus [Chirone and Tornincasa, 2011], [Bertoline et al., 2011], [Lieu and Sorby, 2008], [Hamade et al., 2007], [Allsop, 2009], the following series was proposed to group the main CAD modeling skills or tasks:

1. Drawing profiles.
2. Models obtained by simple extrusion or revolution of profiles.
3. Complex models by combination of different extrusions and revolutions.
4. Models where the reference geometry ("datums") is used to build the structure ("scaffold") of the model (which typically includes oblique elements).
5. Models including curved profiles and sweeps (spherical caps, tori, springs, etc.).
6. Models including replication operations and features.
7. Models of standard parts (Screws, bolts, etc.).

Detailed information on the purpose and contents of every task can be found in [Company and González, 2013].

5. Assertions Maps

Assertions maps are conceived as visual representations of the rubrics where the different assertions of each rubric are displayed along one axis, while the evolution of each assertion throughout the syllabus is displayed along the other axis. The assertions maps are intended to display the expand-contract process.

5.1. Detailed assertions

To quantify their criteria, Amadori et al. [2012] apply complex measurements of the design space, which are clearly not applicable while training novice designers. Rynne and Gaughran [2008] argued that creating robust sketch geometry is the most critical user issue in capturing design intent. They also defined a set of "attributes" of a CAD model, which are similar to the assertions:

- Correct sketch plane selection for base feature sketch
- Optimum model origin
- Correct base feature
- Correct part orientation
- Appropriate use of symmetry planes
- Simple sketch geometry
- Correct sketch relations
- Fully defined sketch geometry
- Correct feature sequence
- Parent-child feature relations
- Correct feature terminations
- Correct feature duplication
- Correct part design intent
- Part accommodates planned and unforeseen design modification without feature failure.

However, these attributes are not formulated to be easily answered in the frame of a rubric. Johnson and Diwakaran [2011] and subsequently, Diwakaran and Johnson [2012] elaborated those attributes, adding detailed descriptions and metrics, most of them binary (true/false).

Working in parallel in CAD Modeling Strategies (CMS), Allsop [2009] proposed a distinction between what she called 'Feature-based', 'Overarching' and 'Detailed' approaches. She used Feature-based instead of the already coined term cognitive modeling, whereas overarching and detailed include some procedural modeling strategies: datum-based modeling (horizontal and skeleton strategies), use of replicating operations (duplicate and symmetry strategies), and actively encouraging the user to consider alternate approaches.

Departing from these attributes and the assertions included in the appendix of [Company et al., 2013], while applying the approach described in previous section, the assertions shown in Table 3.1 are acquired. These assertions constitute one axis of the assertions map.

Table 3.1. Assertions for modeling.

Assertion Code			Description
M1			**The model is valid.**
	M1.1		Model can be retrieved
		M1.1a	Model is not missing (neither lost nor wrongly named).
		M1.1b	There are no errors opening the document of the model.
	M1.2		Model can be used.
		M1.2a	Model tree does not contain error messages.
		M1.2b	Document of the model is compatible with the application.
M2			**The model is complete.**
	M2.1		The model replicates the shape and size of the part.
		M2.1a	The model replicates the shape of the part.
		M2.1b	The model replicates the size of the part.
M3			**The model is consistent.**
	M3.1		Profiles are robust (changes do not produce unexpected failures).
		M3.1a	Profiles do not contain duplicated lines in any modeling operation
		M3.1b	Profiles do not contain segmented lines in any modeling operation.
		M3.1c	Profiles of modeling operations are always fully constrained.
	M3.2		Profiles are flexible (allow many changes)
		M3.2a	Profile constraints are weak enough to allow local changes (design variations).
		M3.2b	Profile constraints are strong enough to prevent local changes (design variations), from causing undesired changes or errors.
	M3.3		Model tree is robust.
		M3.3a	Model is aligned and oriented relative to global reference system.
		M3.3b	Model uses explicit datums to make the scaffold of the shape independent from local shapes and sizes.
	M3.4		Model tree is flexible.
		M3.4a	Functional elements are defined by independent modeling operations.
		M3.4b	The parent/child relations in the model tree do not have needless dependences.
M4			**The model is concise.**
	M4.1		The model does not contain repetitive or fragmented constraints, modeling operations or datums.
		M4.1a	The profiles do not contain repetitive or fragmented constraints.
		M4.1b	The model does not contain repetitive or fragmented modeling operations.
		M4.1c	The model does not contain repetitive or fragmented datums.
	M4.2		Replication operations (translate-and-repeat, rotate-and-repeat and symmetry) are used whenever possible.
M5			**The model is simple (as simple as possible, but not simpler).**
	M5.1		The model tree is clear and understandable.
		M5.1a	Modeling operations are labeled in the modeling tree to emphasize what are they for, instead of how have been built.
		M5.1b	Related modeling operations are grouped in the model tree to emphasize parent-child relationships.
	M5.2		The model uses compatible and standard modeling operations.
		M5.2a	Most compatible modeling operations are always preferred.
		M5.2b	Standard modeling operations are always preferred.
M6			**The model conveys design intent.**
	M6.1		The modeling process is effective in conveying the right information about function.
		M6.1a	The geometric constraints in the profiles help to highlight functional relationships.
		M6.1b	Models use feature-based operations that convey the functionality of the parts.
		M6.1c	Datums convey the skeleton of the model.
	M6.2		The modeling process is efficacious (conveys more design intent than other modeling processes).

	M6.2a	There are no fragmented or overlapped modeling operations.
	M6.2a1	There are no fragmented operations that mask the final part function.
	M6.2a2	There are no overlapped operations that fill what already was filled or empties what already was empty.
	M6.2a3	There are no overlapped operations that mask previous operations.
	M6.2b	There are no deficient feature types, shapes or anchorages.
	M6.2b1	There are no deficient feature types (suitable features are not replaced by merely similar ones).
	M6.2b2	There are no deficient feature shapes.
	M6.2b3	There are no deficient feature anchorages.
M6.3		The model is efficient: it is modeled the most desired way to convey design intent (the same design intent could not be conveyed in a simpler way).
	M6.3a	Modeling operations are labeled in the modeling tree to distinguish between datums, scaffolds, core, detail, replication and cosmetic operations.
	M6.3b	The feature tree reads like a user manual to edit the model.
	M6.3c	Replication operations are used to convey functionality in the model tree.
	M6.3d	Design decisions are traceable within the model tree.

It is noted that the CAD models assertions table is disaggregated to distinguish between robustness and flexibility of profiles and the model tree. The author does not evaluate robustness and flexibility of modeling operations themselves, which, when used in their basic functionality, guarantee those properties. Obviously, more advanced users should evaluate them, as they usually push some operations to their limits.

Assertions related to Dimension 6 were most difficult to determine, because they imply a certain degree of expertise. To this end, negative knowledge is considered, as described by Mandorli and Otto [2013]. They define it as knowing what *not* to do (i.e. "knowing how to avoid grave errors and approaches which are inefficient in certain situations").

The goal was to define assertions that describe actions that induce situations best avoided. So, the author reformulated as assertions those feature deficiencies described by Mandorli and Otto [2013] which were developed as anchor concepts to evaluate CAD models in respect to missing/poor design intent.

One important conclusion derived from Table 3.1 is that some dimensions appear to be opposed to each other. For instance, simple models result from simple and compatible operations (i.e. a negative cylindrical extrusion to produce a drilled hole) while design intent results from complex and sophisticated operations (a drilling feature). The ability to understand these trade-offs between such opposing dimensions and the ability to choose the proper procedure differentiates between novice and expert CAD users. A novice tends to interpret that simple cylindrical holes are better than drills while the expert realizes that conveying design intent is more important, and then subordinates simplicity by choosing the simplest drill available. Assessment strategies are used to introduce such subtle differences.

5.2. Assertions maps

The assertions maps are similar to the time-line for the semester-long "Strategic Use of CAD" course by Bhavnani et al. [1999]. The main difference being that the assertions maps also convey the expand-contract process.

A different rubric is defined for each series. While the six main dimensions are always maintained, the level of detail varies according to the expand-contract criterion:

- The first time a new concept was evaluated, it was described by one or more detailed rubric items.
- In subsequent rubrics, the concept was recursively abstracted.

In the end, every main dimension was directly evaluated, and rubric items were only left as a subsidiary way to allow clarifying the score when the interviewed student felt it was necessary.

Results are shown in Table 3.2. A black background indicates that answering the assertion is mandatory. A grey background specifies that only the higher level assertion is mandatory, while the detailed assertions that complement it are optional. Finally, an "X" means that the assessment cannot continue if this assertion fails. It is a "no-pass" flag.

Table 3.2. Assertions map for models.

	Tasks											
Assertion	1	2	3	4	5	6	7	8	9	10	11	12
M1					X	X	X	X	X	X	X	X
M1.1												
M1.1a												
M1.1b												
M1.2												
M1.2a												
M1.2b												
M2						X	X	X	X	X	X	X
M2.1												
M2.1a												
M2.2b												
M3												
M3.1												
M3.1a												
M3.1b												
M3.1c												
M3.2												
M3.2a												
M3.2b												
M3.3												
M3.3a												
M3.3b												
M3.4												
M3.4a												
M3.4b												
M4												
M4.1												
M4.1a												
M4.1b												
M4.1c												
M4.2												
M5												
M5.1												
M5.1a												
M5.1b												
M5.2												
M5.2a												
M5.2b												
M6												
M6.1												
M6.1a												
M6.1b												
M6.1c												
M6.2												
M6.2a												
M6.2a1												
M6.2a2												
M6.2a3												
M6.2b												
M6.2b1												
M6.2b2												
M6.2b3												
M6.3												
M6.3a												
M6.3b												
M6.3c												
M6.3d												

As shown in Table 3.2, while solving the first task, CAD trainees are taught about missing electronic documents and making mistakes that result in error messages in the model tree. Then, they are evaluated on Dimension 1, by way of four detailed assertions. At the end of the second task, they are evaluated once more. But this time, some assertions are contracted (i.e. M1.1 is initially expanded into M1.1a and M1.1b2, and is later contracted back to M1.1). At the end of the third task, validity is globally evaluated; although, to assist them to fix the concept, the detailed assertions are also provided. At the end of the fourth task, only the global dimension of validity is evaluated. While assessing further tasks, validity is not given a numeric score. On the contrary, lack of validity is used as a no-passing criterion.

Consistency is progressively introduced during Sessions 1 to 4, and is compacted starting from the fifth session. Conciseness is progressively introduced during Sessions 3 to 6, and is condensed starting from the seventh session.

Simplicity and capture of design intent require greater levels of expertise and are concepts of increasingly abstract nature; hence, they require more time to be comprehended. Sessions 9 to 12 (which are primarily concerned with introducing drawing extraction and assemblies) are used to reinforce such dimensions. Our assumption in justifying more than seven sessions in the sequence is that since models interact with assemblies and drawings, it is not until CAD trainees attempt to produce assemblies and drawings that they realize that some subtle mistakes or bad practices are embedded in their models.

The six quality dimensions are divided into three groups: a) nearly dichotomous, b) require internal arrangement, and c) require external trade-offs. In principle, validity and completeness are dichotomous dimensions, as they either are or are not accomplished, so they could be easily controlled though a checklist (as proposed by Gebhard [2013]). Replacing dichotomous checklists by five-point Likert scales during training periods assists teachers to stimulate the development of their students by introducing intermediate scoring values to visualize their progress. Once the dimensions have been understood, they are simplified as purely dichotomous.

Consistency belongs to the second category, as it clearly requires an internal compromise between robustness and flexibility. It must be this way because designers require models simultaneously robust and flexible in order to explore new solutions. An excessively robust design prevents creativity, while a too loose design behaves erratically, even when undergoing simple changes. The ability to understand and solve these trade-offs is a main skill in engineering design. Rubrics make them visible and it is argued that explicit procedures conveyed through suitable teaching provide students with effective coping strategies.

Conciseness and simplicity belong to the third category. Conciseness may be seen as a dichotomous dimension if isolated, but using the higher level modeling operations easily compromises simplicity. So, both dimensions are linked through what can be called external trade-offs. Obviously, the sixth dimension also belongs to the third category, as it interacts in the same manner with Dimensions 4 and 5. These mutual and contradictory interdependencies suggest longer maturing times for those three dimensions in the expand-contract map.

It must be highlighted that the individual assertions included in Table 3.1 cannot be understood as categorical goals, but are only simplifications. For instance, it is quite obvious to expert designers that completeness depends on future use. However, the detailed assertions shown in M2 guide novice students to produce full models (including their details). It is only when assertion M6.3a is used (Task 9 in Table 3.2), that the author begins to distinguish between "core, detail, replication and cosmetic operations." In this stage of the training period, it is suggested that detail and cosmetic operations must be introduced, explaining their advantages (i.e. they allow for more realistic models) and disadvantages (they incur too much calculation time without noticeably improving the accuracy of CAE analysis). In other words, if (1) the primary view includes a full model and (2) its model tree is ordered from core to detail, then it is always easy to obtain simplified secondary views without a loss

of design intent. However, beginning students cannot cope with such a complex strategy, so they must be directed to discover this concept gradually, which is the intention of an assertions map.

In a similar way, it is well known that under-constraining is an effective strategy in several cases. Learning to under-constrain is difficult, as it requires an appreciation of the process of balancing robustness and flexibility, which is why this concept evolves along the training period. Initially, novice students are trained to always constrain their models, but only after they have consolidated the habit of constraining their models can they be safely informed about the exceptions. Firstly, because only then are they ready to understand them as exceptions. Secondly, since understanding the potential benefits of under-constrained models makes them also conscious of the potential dangers implied in a careless use of such a strategy. This awareness is what the expand-contract strategy attempts to provide: the capability of first putting the focus on simple and isolated criteria while progressively introducing more sophisticated mutual relations and agreements between opposite conditions. For instance, the first time students are exposed to the third dimension of the rubric, they are asked to evaluate three simple and supposedly independent criteria in order to measure the robustness of profiles: M3.1a, M3.1b and M3.1c. In parallel, they are asked to evaluate the flexibility of profiles through two simple criteria: M3.2a and M3.2b. Later, they are asked to evaluate the robustness of profiles as a single item (M3.1), which requires agreement between all three previous criteria. Flexibility of profiles is condensed in a similar fashion. In other words: initially, fully constraining the profiles is mandatory, while in the second stage, fully constraining the profiles only helps to ensure that changes do not produce unexpected failures, where "unexpected failures" is a qualitative concept requiring agreement between robustness and flexibility. In the third and final stage, students are asked to acquire "consistent" models, which result from balancing robustness and flexibility, at both the levels of profiles and model trees.

6. More Lessons Learned on Using Rubrics

Some previous lessons learned in [Company et al., 2014] led to the development of the assertion map detailed in Table 3.2, but some others were found to be undeveloped when guiding decisions about the design and implementation process of quality oriented CAD rubrics. Hence, three pilot experiments were performed to gain knowledge on those topics. Three main aspects are described next: if rubrics may be aggregated, what is the right speed for the expand-contract process, and what is the best method to convey procedures and tools that help in measuring these different assertions. The experiments are explained in detail to illustrate the suggested process for replicating this approach for further developing rubrics adapted to other scenarios.

6.1. Interspersed Rubrics

One important lesson learned in [Company et al., 2014] is that particular and specific rubrics are the only valid choice to share and convey quality criteria. Obviously, the assertions map assumes an incremental process in which assertions are introduced sequentially during the training period of novice users. However, nothing was concluded about whether rubrics must be separated or may be interspersed in the same form during the actual evaluation processes. A new experiment was conducted aimed at determining whether each task requires a separate rubric. Students were asked to solve and evaluate three tasks: A) obtain the solid model of the casing with flaps depicted in Figure 3.1; B) assemble previously modeled parts of a shut-off valve (Figure 3.2), and C) obtain the assembly drawing of the valve, including part numbers and bill of materials.

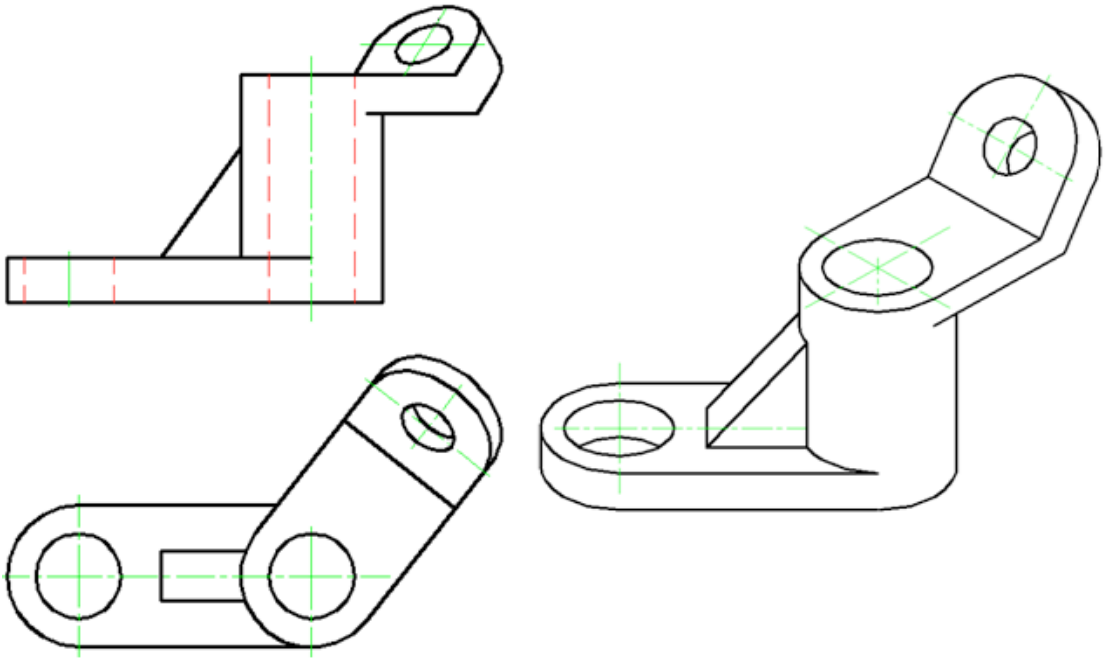


Figure 3.1. Casing with flaps.

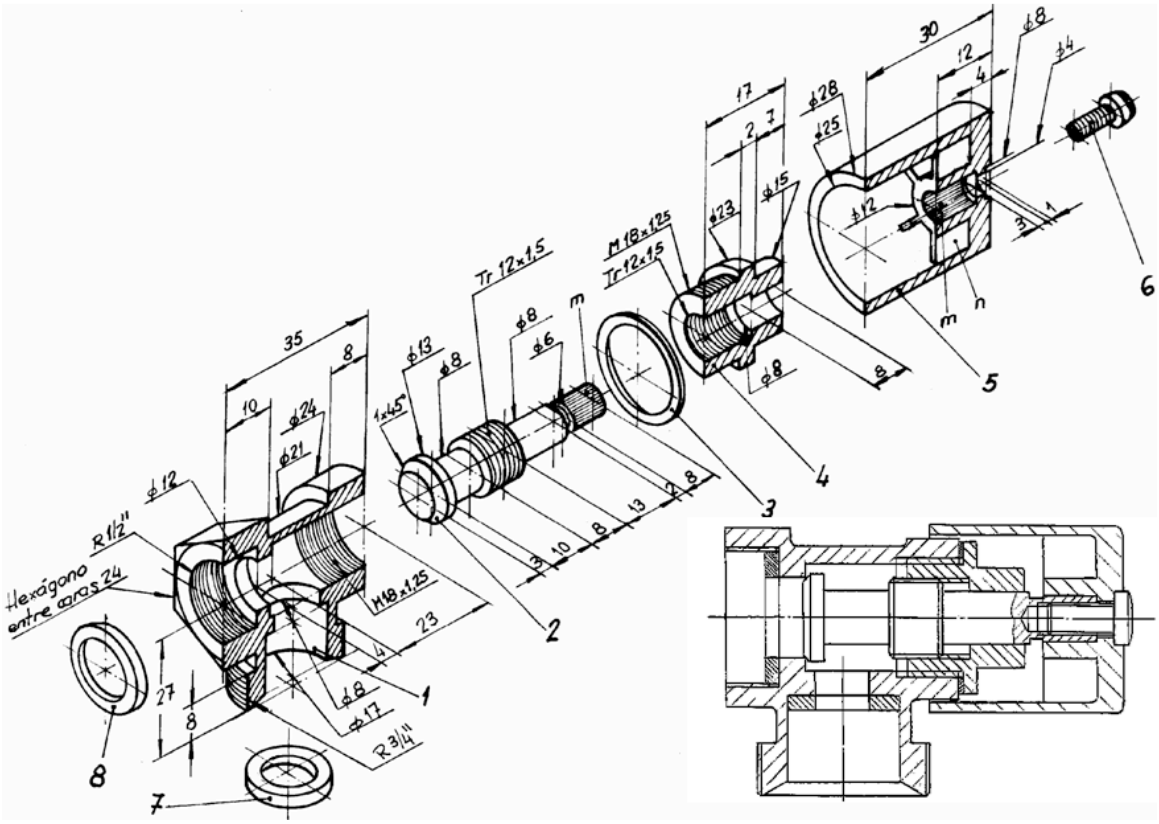


Figure 3.2. Shut-off valve.

The form that contains all the assertions used to evaluate the three tasks is shown in Figure 3.3 (shown above). It differs from Table 3.1 because the experiment was completed previous to solidifying the final set of assertions detailed in the table. Also, it was split into two forms: one for Task A and a second rubric with assertions corresponding to Tasks B and C interspersed. The students were asked to evaluate whether they felt that each rubric was helping them to understand and solve the tasks,

according the five-point Likert scale. Their average answer was 0.68 (quantified in the range [Strongly disagree= 0, Strongly agree= 1]), which means the value is closer to agree than to neither agree nor disagree. This moderately optimistic reply contradicts the fact that they found the rubrics difficult to use, since of the 16 interviewed students, five of them (31%) returned incomplete rubrics. The difference between the first form (Task A) and the second form (Tasks B and C) was moderate and better for the second (0.63-0.73). So, it is established that this information does not help to conclude whether separate forms are better than interspersed ones. Discrepancies between self-evaluation and teacher evaluation for Task A (Table 3.3) and for Tasks B and C (Table 3.4) were also compared.

Item	Assertion	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Do not understand the assertion
1	**The project documentation is valid**						
1.1	There are no errors opening documents						
1.2	Models do not contain error messages						
1.3	Assemblies do not contain error messages						
1.4	Drawings do not contain error messages						
2	**The project documentation is complete**						
2.1	All non-standard parts have been modelled						
2.2	All of the models replicate the shape and size of their respective parts						
2.3	Assemblies include all of their parts, and only them						
2.4	All of the parts are suitably placed in their assemblies						
2.5	Assemblies replicate mating conditions of the products they represent						
2.6	All drawings (excluded those of standard parts) have been included						
3	**The project documentation is consistent**						
3.1	Models allow local changes (design variations), without causing undesired changes or errors						
3.2	The profiles of any generalized sweep operations are always fully constrained						
3.3	Assemblies allow valid movements, without causing undesired movements						
3.4	There are no missing views or cuts in the drawings						
3.5	There are no missing "symbols" in the drawings (dimensions, detail numbers, bill of materials, manufacturing symbols, etc.)						
3.6	Drawings meet the standards						
4	**The project documentation is concise**						
4.1	The models could not be obtained with a substantial reduction of operations						
4.2	The assemblies could not be assembled with a substantial reduction of mating conditions						
4.3	Replication pattern operations (symmetry, matrix) are used whenever possible						
4.4	There are none or few redundant views and cuts in the drawings						
4.5	There are no unnecessary symbols in the drawings (dimensions, detail numbers, bill of materials, manufacturing symbols, etc.)						
5	**The project documentation conveys design intent effectively**						
5.1	The constraints in the profiles meet the design intent						
5.2	Models use all suitable datums						
5.3	Models do not use unnecessary datums						
5.4	Models use design features						
5.5	Assemblies use mating conditions which highlight design intent instead of geometry						
5.6	Assemblies are subdivided into sub-assemblies which convey functionality						
5.7	Links between assemblies and models are used to automate bills of materials and detail numbers						
5.8	Drawings are mainly extracted from models or assemblies (i.e. contain none, or very few, manually added "cosmetic" elements)						
6	**The project documentation is clear and comprehensible**						
6.1	Modelling operations are labelled to convey design intent in the modelling tree						
6.2	Mating conditions are labelled to convey design intent in the assembly tree						
6.3	Sheet formats and scales are suitable for the project						
6.4	Documents are well organized (names of files and structure of folders)						

Figure 3.3. Full rubric with interspersed assertion on modelling, assembling and extracting drawings.

Table 3.3. Differences between the self-evaluation and the teacher evaluation in Experiment 1, Task A (void means no difference; 1 means 100% difference).

Student	1.1	1.2	2.2.1	2.2.2	3.1	3.2	4.1	5.1.2	5.2	5.3	5.4	6.1	|Avg| A
#1	0.25	0.25	0.25	0.5		0.25	0.25	0.25	0.5		0.5	0.75	0.31
#3	0.25	0.5			0.25	0.25	0.25		0.25		0.75	0.5	0.25
#4			0.25	0.25		0.25	0.25		0.25		0.25	1	0.21
#6		0.25	0.25	0.25	0.5			0.75	0.25		1	0.75	0.33
#7				0.25	0.25	0.25	0.5				0.5	0.75	0.21
#9	0.25	0.25		0.25	0.25	0.25	0.25	0.25	0.75	0.5	0.5	0.25	0.31
#10	0.25	0.25	0.25		0.5	0.25	0.25		0.25	0.5	0.5	0.75	0.31
#12	0.25	0.5		0.25	0.25	0.25		0.5	0.25	0.25	0.5	1	0.33
#14	0.25			0.25	0.25		0.5	0.25			0.25	0.75	0.21
#15		0.25	0.5	0.75	0.5		0.25	0.5	0.25	0.5		0.5	0.33
#16	1	0.75	0.75	0.25			0.75	0.5	0.5	0.25	0.75	0.25	0.48
|Avg|	0.23	0.27	0.20	0.27	0.25	0.16	0.30	0.27	0.30	0.18	0.50	0.66	0.30

Table 3.4. Differences between the self-evaluation and the teacher evaluation in Experiment 1, Tasks B and C (void means no difference; 1 means 100% difference).

Student	1.1	1.3	1.4	2.3	2.4	2.5	2.6	3.3	3.6	4.2	4.4	4.5	5.5	5.6	5.7	5.8	6.2	6.3	|Avg| B	|Avg| C
#1	1		0.25		0.5	0.25	0.5	0.25	0.75		0.75	0.25	0.5	0.5	0.5		0.75	0.25	0.38	0.41
#3					0.25	0.25	1		0.75	0.5	0.25	0.5	0.25	0.5	0.75		0.75	0.25	0.25	0.44
#4				0.25		0.25	0.25	0.25	0.75		0.25	0.25	0.5	0.75	0.25	0.25	0.75	0.5	0.28	0.31
#6		0.25			0.75	0.75		0.5	1	0.25	0.25	0.75	1	0.75	0.25		1	1	**0.53**	0.41
#7				0.25	0.25	0.25	0.25	0.25	0.25				0.5	0.75	1		0.75	0.5	0.30	0.25
#9	0.25	0.25	0.25	0.25		0.25				0.25		0.25	0.75	0.75	0.75	0.25	0.75	0.25	0.35	0.22
#10								0.25			0.25	0.25	0.25	0.5			1	0.5	0.20	0.13
#12	0.25	0.25	0.5		0.25	0.25	0.25		0.25	0.25	0.5	0.5	0.5	0.75	0.25	0.25	0.75	0.25	0.33	0.34
#14	0.25		0.25	0.25	0.25	0.25	0.25		0.75	0.25	0.75	0.25	0.25	1	0.25		0.75	0.75	0.33	0.41
#15	0.5	0.25	0.25	0.25	0.25	0.5	0.25		0.5	0.25	0.25	0.25	0.25	0.5	0.25	0.25	0.75		0.35	0.25
#16	0.5	0.5	0.75	0.25		0.25	0.5	0.25	0.25	0.25	0.25		0.75	0.75	0.25	0.25	0.75	0.5	0.43	0.34
|Avg|	0.25	0.14	0.20	0.14	0.23	0.30	0.30	0.16	0.48	0.18	0.32	0.30	0.50	0.68	0.41	0.11	0.80	0.43	0.34	0.32

In viewing the tables, it concluded that there are no significant differences. Hence, the pilot experiment does not allow the making of any strong hypothesis. It cannot be concluded whether separate are better than interspersed forms. The only clues in favor of independent forms are the high ratio of students that failed in returning the rubrics (31%) and the subjective evaluation of the teachers that reported that marking interspersed forms was more time consuming and prone to error. However, this is still an open problem, and a full size experiment including a control group is clearly required to validate or reject this hypothesis.

6.2. Expand-contract process

The goal in this experiment was determining whether the expand-contract process was necessary. The question posed was: do students understand dimensions at first sight, or do they need time and intermediate stages to become acclimatized to the main dimensions?

Rubrics were developed following the assertions map and undergraduate students were "exposed" to them. Exposed means that they were informed about their existence, were instructed where to find them (in the virtual classroom site, which is a customized Moodle website), and were generically taught during the practical classes about the meaning and importance of these rubrics. However, they were not required to fill out any of them before the experiment, neither were they specifically shown how to interpret their assertions or measure their degree of accomplishment.

In parallel, the same subject was offered to first year Master's students. This group of post-graduate students was primarily in mechanical, design, and other engineering disciplines. They followed the same program as the undergraduate students [Company and González, 2013]. In principle, course content should be repetitive, but since they possess different backgrounds, homogenization was required. Furthermore, the teaching strategy was varied by reducing theoretical introductions, emphasizing more advanced aspects, and requiring them to solve more complex problems (that is, the last examples in each series).

Both groups were required to solve the same exam problem. Figure 3.4 shows a plughole created as a cut view. Its scale may be fixed though a single given dimension and the bill of materials was also listed. Figure 3.5 represents the detail drawing of Part 3, which is constructed of brass (Item 3). Students were provided complementary information: the upper body of the plughole includes 8 holes to allow water flow; the lower plughole body has 4 union arms between the central anchoring element and the outer funnel, and all threads are ISO metric. In the assembly, the separation between Parts 1 and 3 is due to the (undefined) space occupied by kitchen sink: Parts 1, 2 and 4 are assembled from the upper side of the sink hole; the other parts are assembled to the drain pipe, located below the sink hole (and not identified with a part number in Figure 3.4). Solid models had been obtained for all the non-standard parts but 3, and were provided during the exam. Students were asked to: A) model the lower plughole body (detail Number 3), as detailed in Figure 3.5; B) create the drawing of Part 1 (the drawing was to be obtained by extracting it from the solid model given by the instructor), and C) create the assembly. Students could freely use the models provided during the exam (Parts 1, 2, 5, 6, and drain pipe), together with the required models of the standard parts obtained from the library (Solidworks Toolbox®).

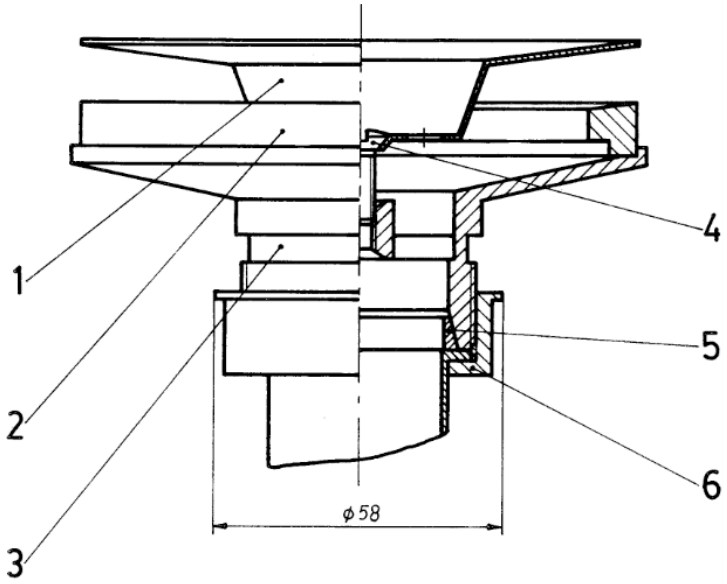


Figure 3.4. Plughole assembly.

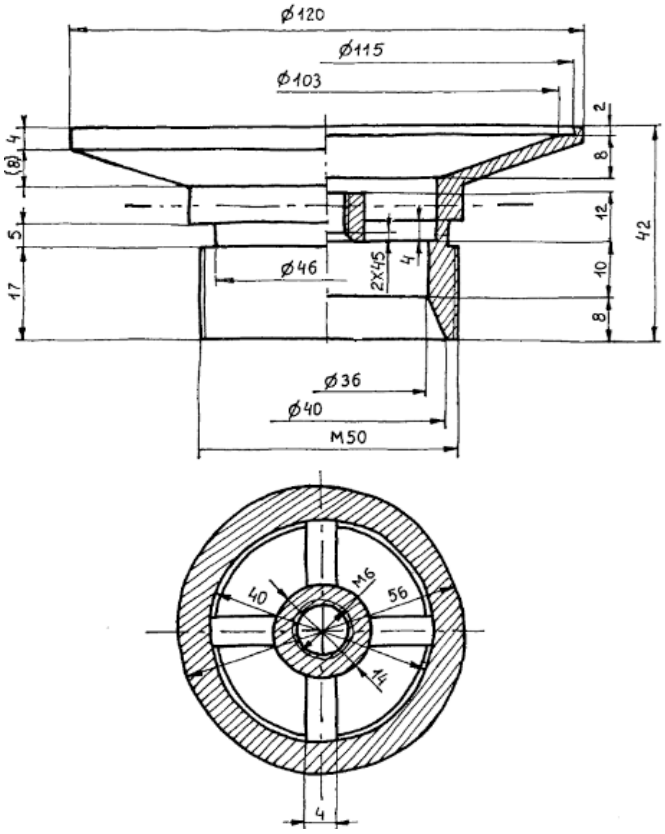


Figure 3.5. Lower plughole body.

In order to familiarize students with rubrics, Task D explicitly required self-evaluation, and a reward was given to those matching the teacher evaluation.

40 students were examined, and received 30 valid rubrics were submitted. Twelve of them (40%) always marked only the main dimensions. Four students (13%) always marked the main dimensions and periodically (when they thought it was opportune) also marked some auxiliary assertions. Eight students (27%) marked all the assertions in all three rubrics while four students (13%) only marked detailed assertions. Finally, two students failed to follow any consistent answering pattern.

The main result is that direct "immersion" into contracted rubrics does not work. Neither the undergraduate nor the Master's students obtained good results. Only 53% of the students completed the form in the condensed way, while the remaining students used the expanded version of the form. It is quite clear that some "simple" assertions are well understood by most students, but other assertions still require training in order to be understood. Hence, the expand-contract process is clearly beneficial, but it should be evaluated in a full-scale experiment aimed at determining the precise contraction speed for each assertion.

The experiment provided other valuable information and these results are presented in Tables 3.5 and 3.6. In both tables, the three tasks are tabulated separately. In addition, each task has been labeled to indicate whether main assertions were marked (Yes or Void), and whether auxiliary assertions where marked (Yes, Sometimes, or Void).

Table 3.5. Rubrics of undergraduate students (Green cells indicate that the subject marked main assertions, orange indicates they also marked auxiliary assertions).

Subject	1	2	3	4	5	6	Task A	Main	Auxiliar	1	2	3	4	5	6	Task B	Main	Auxiliar	1	2	3	4	5	6	Task C	Main	Auxiliar
1		0.3	0.3	0.5		0.5	0.25	Y	S	0.3		0.3	0.3		0.3	0.17	Y		0.8	0.3	0.3		0.5	0.5	0.38	Y	
2		0.3	0.5	0.8	0.5	0.8	0.46	Y		1		0.3	0.3	0.8	0.8	0.50	Y		1	0.5	1		0.5	0.5	0.58	Y	
3	0.3	0.4	0.3		0.6	0.6	0.36	Y	Y	0.4		0.3		0.3	0.3	0.22	Y	Y	0.8		0.3	0.5	0.3	0.5	0.40	Y	Y
4			0.3		0.3	0.5	0.17	Y		0.8	0.3	0.3	0.3	1	0.8	0.54	Y		0.8	0.3	0.3	0.5	0.3	0.3	0.38	Y	
7		0.3	0.3	0.3	0.8	0.8	0.38	Y		0.8	0.4	0.5	0.8	0.8	0.8	0.65	Y	S		0.3	0.3		0.8	0.3	0.25	Y	S
8	0.5	0.3	0.3	0.3	0.5	0.8	0.42	Y		0.5		0.3	0.3	0.5	0.8	0.38	Y				0.5		0.3	0.3	0.17	Y	
9		0.3	0.5	0.3	0.5	0.8	0.38	Y			0.3	0.5	1	0.5	1	0.54	Y			0.3			0.5	0.5	0.21	Y	
10	1	0.8	0.5	0.3	0.3	0.3	0.50	Y		1		0.3	0.3	0.3		0.29	Y		1	0.8	0.5	0.8	0.5	0.5	0.67	Y	
12	0.3			0.3		0.3	0.13	Y							0.5	0.08	Y		0.3	0.5	0.3	0.3	0.3	0.3	0.29	Y	
13					0.3		0.09	Y	Y	0.4		0.3	0.3	0.3	0.3	0.24	Y	Y					0.6	0.5	0.25	Y	Y
14	0.3		0.4	0.8	0.6	0.4	0.43		Y	0.3		0.3		0.7	0.5	0.28		Y	0.3		0.4		0.5	0.8	0.31		Y
16	0.8		0.4	0.6	0.6	0.6	0.50		Y	0.3				0.3	0.3	0.13		Y	0.8	0.6		0.5	0.5	0.4	0.46		Y
17	0.5			0.3	0.3	0.8	0.29	Y			0.5	0.5	1	0.5	1	0.58	Y		0.5	0.3	0.3		0.5	0.5	0.35	Y	S
18							0.06	Y	Y	1	0.9	1	0.8	0.8	1	0.90	Y	S			0.3	0.3	0.3	0.5	0.21	Y	
19	0.3	0.8	0.3	0.3	0.5	0.5	0.42	Y			0.3	0.5	0.5	0.5	0.3	0.33	Y		0.8	0.5	0.3	0.8	0.8	1	0.67	Y	
21		0.3		0.3	0.5	0.4	0.25	Y	Y		0.3	0.8		0.6	1	0.44	Y	Y				0.3	0.7	0.6	0.27	Y	Y
22					0.5	0.5	0.17	Y		0.3						0.04	Y		0.8	0.8	0.5	0.5	0.5	0.8	0.63	Y	
23	0.3	0.3	0.3	0.3		0.3	0.21	Y		0.8		0.5	0.5	0.3	0.5	0.42	Y		0.3		0.3	0.3	0.3	0.3	0.21	Y	
24					0.5	0.3	0.17	Y	Y						0.8	0.15	Y	Y				0.3	0.3	0.3	0.17	Y	Y
25	0.8		0.4	0.5	0.3	0.4	0.41		Y	0.5				0.4	0.8	0.34		Y	0.5	0.3		0.3	0.4		0.28		Y
26	0.3	0.3		0.3	0.5	0.3	0.25	Y			0.3	0.8	0.3	0.3	0.8	0.38	Y		0.3	0.3	0.3	0.3	0.3	0.3	0.25	Y	
27	0.3	0.3	0.3	0.3	0.3	0.3	0.25	Y		0.3				0.3	0.8	0.21	Y			0.3			0.5	0.5	0.21	Y	
29				0.3	0.4	0.3	0.20	Y	Y					0.3	0.5	0.19	Y	Y			0.3	0.4	0.3	0.3	0.22	Y	Y
30	0.4		0.4			0.4	0.20	Y	Y	0.3	0.5	0.5	0.3	0.8	0.8	0.50	Y	Y	0.4			0.3	0.5	0.4	0.28	Y	Y
32		0.3				0.4	0.16	Y	Y	0.3	0.3	0.3	0.3	0.7	0.8	0.40	Y	Y				0.3	0.5		0.18	Y	Y
33	0.3	0.3	0.3	0.4	0.8	0.5	0.40	Y	S		0.3		0.3	0.5	0.5	0.25	Y		0.3	0.5	0.8	0.8	0.8	0.8	0.63	Y	
34	0.3	0.5	0.5	0.8	0.5	0.8	0.54	Y		0.5		0.3	0.5	0.5	0.8	0.42	Y		0.5	0.3	0.5	0.5	0.8	0.8	0.54	Y	
36			0.7	1	1	0.8	0.60	Y	Y		0.3		0.3	0.3	0.5	0.24	Y	Y					0.5	0.8	0.23	Y	Y
39		0.3	0.4	0.8	0.8	0.5	0.44		Y		0.3	0.4	0.3	0.5	0.8	0.35		Y			0.4	0.3	0.5		0.20		Y
40	0.3				0.6	0.3	0.20	S	Y	0.3	0.3	0.4		0.5	0.5	0.33		Y	0.3				0.5	0.4	0.22		Y
Avgl	0.22	0.20	0.26	0.30	0.41	0.46	0.31			0.32	0.18	0.30	0.28	0.43	0.53	0.35			0.33	0.23	0.27	0.26	0.47	0.45	0.34		

Table 3.6. Rubrics of Master's students (Green cells indicate that the subject marked main assertions, orange indicates they also marked auxiliary assertions).

Subject	1	2	3	4	5	6	Task A	Main	Auxiliar	1	2	3	4	5	6	Task B	Main	Auxiliar	1	2	3	4	5	6	Task C	Main	Auxiliar
1				0.3			0.10	Y	Y				0.3	0.5	0.5	0.22	Y	Y			0.3	0.3	0.8	0.8	0.35	Y	Y
2				0.5			0.08	Y	Y			0.5	1	0.5		0.33	Y					0.3	0.8	0.5	0.25	Y	S
4	0.3	0.3	0.3	0.3		0.8	0.29	Y		0.3	0.3	0.3	0.3	0.5	0.5	0.33	Y		0.5		0.3	0.3	0.5	0.5	0.33	Y	
5					0.3	0.4	0.13	Y	Y		0.5		0.4	0.7	0.8	0.41	Y	Y	0.3		0.3	0.3	0.5		0.25	Y	Y
6				0.3		0.3	0.09	Y	Y		0.3	0.3	0.3	0.5	0.3	0.25	Y		0.3	0.4	0.3		0.3		0.19	Y	S
7		0.3			0.5	0.6	0.25	Y	Y	0.5	0.3		0.3	0.3	0.8	0.33	Y			0.5	0.3	0.3	0.5	0.5	0.33	Y	Y
8	0.3	0.3	0.5	0.5	0.8	1	0.54	Y		0.3	0.3	0.3	0.3	0.3	0.3	0.25	Y		0.3	0.3	0.3	0.3	0.3	0.3	0.25	Y	
9	0.3	0.4					0.13	Y	Y	0.3	0.3	0.4		0.4	0.3	0.26	Y	Y	0.3		0.3	0.4			0.19	Y	Y
10				0.3		0.3	0.09	Y	Y	0.3	0.3	0.8		0.7	1	0.49	Y				0.3		0.8	0.3	0.21	Y	
11			0.3		0.3		0.08	Y	S		0.3		0.3		0.8	0.21	Y		0.3		0.3		0.3	0.3	0.17	Y	
12		0.4	0.5	0.3			0.19	Y	S		0.3				1	0.24	Y	Y			0.3	0.3	0.3	0.3	0.17	Y	Y
13		0.5	0.8	0.3	0.5	0.5	0.42	Y					0.5	0.5	0.8	0.29	Y	S	0.3	0.3	0.3	0.3	0.3	0.5	0.29	Y	
14	0.3		0.3	0.3	0.3	0.8	0.33	Y	S	0.3		0.3	0.3	0.5	0.8	0.33	Y		0.3	0.3		0.3	0.5	0.5	0.29	Y	S
15		0.3		0.5	0.4	0.3	0.23	Y	Y					0.4	0.3	0.13	Y	Y			0.5	0.3	0.4	0.3	0.24		
Avgl	0.07	0.17	0.22	0.24	0.21	0.35	0.21			0.13	0.18	0.21	0.27	0.40	0.55	0.29			0.16	0.14	0.23	0.21	0.43	0.33	0.25		

In addition, it is possible to qualitatively perceive two main facts:

1. Master's students performed at a higher level. Their evaluations are more accurate than those of the undergraduate students. This performance difference is not due to their background in rubrics, as both groups of students have the same background. The only clear difference is the background in 3D CAD, as many Master's students had been previously exposed to some training in CAD 3D (although not necessarily with SolidWorks).

2. Students appear to understand the rubric used to evaluate Task A, and (to some extent) Task C, but fail in understanding the rubric aimed at evaluating Task B. All of the students had been exposed to rubrics aimed at evaluating 3D models, but none had previous experience with rubrics aimed at evaluating drawings and assemblies. Hence, previous experience with rubrics seems to make a perceptible difference.

It also can be perceived that Dimensions 5 and 6 of Task B are scarcely understood, and, in general, Dimensions 5 and 6 are poorly understood.

Apart from the tabulated results that originate from the students, there is also the instructors' assessment. Instructors do not feel comfortable marking Task B: Dimensions 5 and 6 are confusing (a clear criterion to mark the exams is missing). To overcome this difficulty, teachers verified whether their global perception matched the average qualification for each task. In other words, they adjusted the local qualifications to match a global perception. It is also clearly perceived that some students appear to achieve a passing score by following an "average criterion": these students assign the identical average evaluation to all the dimensions. Presumably, this average has been qualitatively obtained beforehand. Students should be interviewed to validate or reject this perception.

Some students attempt to mark average values for dimensions after first marking their auxiliary assertions, but are unsuccessful because the discrete form forces round-offs (does not allow for continuous average values). Forms based on computer documents could assist to automatically calculate the continuous average value, which would prevent incoherencies between the evaluation of main dimensions and their subordinated assertions. It would also help the students to understand the impact of every evaluation in the total.

Finally, some students partially mark the assertions and other times mark the global value for the dimension. In such cases, if both marks are inconsistent, it is difficult to determine the intended average value for the dimension. Again, an electronic form that automatically calculates the average for each dimension and the final mark would be helpful.
The first group of 40 undergraduate students passed a subsequent test. At the end of the training period, they were asked to model and self-evaluate the six non-standard parts belonging to the same assembly from [Chirone and Tornincasa, 2011] (shown in Figure 3.6).

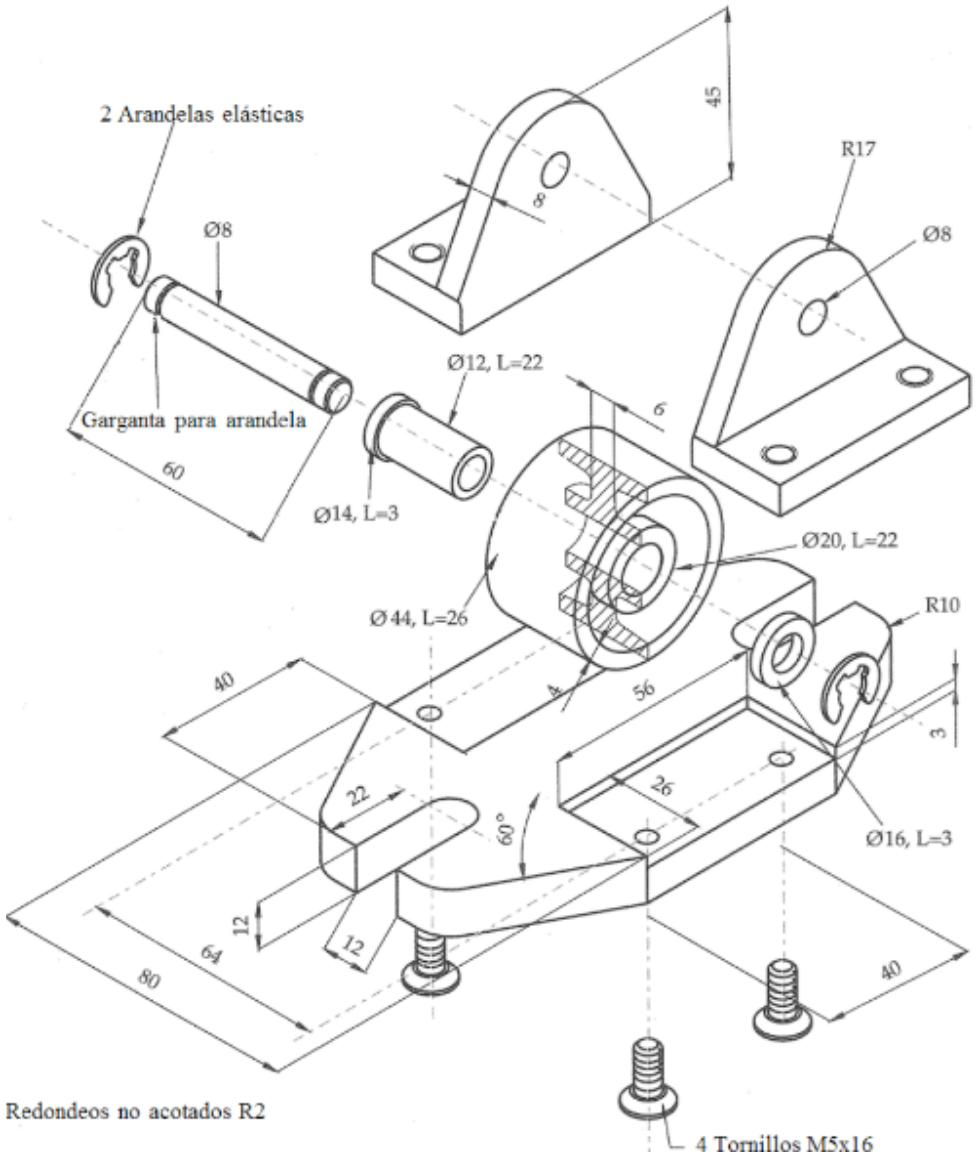


Figure 3.6. Pulley.

Five out of the forty students returned non-valid rubrics (incomplete or containing mistakes). The differences between the self-evaluation and the teacher evaluation for the remaining 35 students are tabulated in Table 3.7. In analyzing the average differences (last row in the table), it can be easily concluded that the first dimension is well understood in all cases, but understanding the remaining dimensions clearly depends on the complexity of the part. For simple parts (i.e. the spacing washer), all but the last dimension have been understood, while complex parts still require additional training before contracting the dimensions. It is of particular note that an apparently simple part like the axle produced important modeling differences because most students didn't use patterns and symmetry to ensure that both slots were equal and symmetrically located. Some students also failed in modeling chamfers as separate operations.

The second conclusion of the experiment relates to cases where non-valid or non-complete models were delivered by the students (cells highlighted in yellow). Following the assertions map, those two dimensions had been labeled as no-pass criteria. Students that delivered non-valid or non-complete

parts (4 out of 40, i.e. 10%), were also unable to detect their mistakes and to suitably mark the corresponding rubrics.

Table 3.7. Condensed rubrics of the six parts of the pulley.

Subject	1	2	3	4	5	6	Base	1	2	3	4	5	6	L bracket	1	2	3	4	5	6	Axle	1	2	3	4	5	6	Shaft sleeve	1	2	3	4	5	6	Wheel	1	2	3	4	5	6	Spacing washer
1			0.25		0.25		0.08							0.00					0.25		0.04					0.50	0.35	0.24				0.50	0.25		0.13							0.00
2	0.25	0.25	0.50	0.25		0.50	0.29	0.75	0.50	0.50	0.50	0.50	0.50	0.54	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.25	0.25	0.50	0.25	0.25	0.75	0.38	0.25	0.25	0.50	0.25	0.25	0.75	0.38	0.25	0.25	0.50	0.25	0.25	0.75	0.38
3	0.25	0.25	0.25	0.25		0.50	0.25	0.25			0.25	0.25	0.50	0.21	0.25	0.50	0.25	0.75	0.75	0.75	0.54	0.50	0.25	0.25	0.50	0.75	0.50	0.46	0.50	0.25			0.25	0.50	0.25	0.25	0.25	0.25	0.25	0.25	0.75	0.33
4					0.75	0.50	0.21	0.25		0.50	0.50	0.50	0.50	0.38	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.75	0.75	0.50	0.50	0.75	0.56	0.25	0.25	0.25	0.25	0.25	0.75	0.33
5	0.25		0.50	0.75	0.75	0.75	0.50				0.25	0.25	1.00	0.25				1.00	1.00	1.00	0.50						1.00	0.17		0.50		0.25	0.25	1.00	0.33						1.00	0.17
6	0.25					0.25	0.08	0.25	0.25				0.25	0.13	0.25	0.25	0.25	0.25	0.25	0.25	0.25				0.50		0.65	0.19						0.25	0.04						1.00	0.17
7					0.25		0.04		0.25					0.04			0.25				0.04						1.00	0.17						1.00	0.17						0.50	0.08
8				0.25	0.25		0.08						0.25	0.04							0.00				0.50	0.25	0.65	0.23		0.50		0.50	0.50	0.50	0.33							0.00
9						1.00	0.17		0.25				1.00	0.21		0.50	0.75	0.50	0.50	1.00	0.54						1.00	0.17		0.50		0.50	0.50	1.00	0.42						1.00	0.17
11	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.25		0.50	0.25	0.25	0.75	0.33	0.25		0.25		0.50	0.75	0.29	0.25	0.25	0.25	0.75	0.75	0.75	0.50	0.25	0.25	0.25	0.75		0.75	0.38				1.00	1.00	1.00	0.50
12	0.25		0.50	0.50	0.25		0.25	0.25			1.00	0.25	0.75	0.38		0.75	0.50	1.00	1.00	0.50	0.63	0.25	0.75		0.75	1.00	0.75	0.58	0.25	0.50		0.50	0.25	0.75	0.38	0.25					0.75	0.17
13						0.25	0.04	0.50		0.25	0.25	0.50	0.25	0.29	0.25		0.50	0.25		0.25	0.21	0.25	0.25	0.25	0.50	0.50	0.65	0.40	0.25	0.50	0.25	0.25	0.50	0.25	0.33	0.50	0.50	0.50	0.50	0.50	0.50	0.50
14							0.00		0.25					0.04		0.25	0.25				0.08						0.30	0.15							0.00							0.00
15						0.50	0.08		0.25					0.04		0.75	0.50	1.00	1.00		0.54				1.00	1.00		0.33		0.50		0.50	0.50		0.25							0.00
16						0.25	0.04						0.25	0.04						0.25	0.04						0.65	0.11						0.25	0.04					0.50	0.25	0.13
17					0.25	0.25	0.08		0.25			0.25	0.25	0.13				0.75		0.25	0.21				0.50	0.25	0.65	0.23				0.25	0.25	0.25	0.08					0.25	0.25	0.08
18		0.25	0.25	0.25	0.25	0.25	0.21	0.25	0.25	0.75	0.25	0.50	0.25	0.38	0.50		0.25		0.25		0.17	0.50	0.25	0.25	0.25	0.25	0.15	0.28	0.25	0.25	0.50	0.50	0.50	0.25	0.38	0.25	0.25	0.25	0.25	0.25	0.75	0.33
19	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.75	0.25	0.25	0.33	0.25	0.25	0.25	0.25	0.25	0.65	0.32	0.25	0.25	0.25	0.50	0.50	0.25	0.33	0.25	0.25	0.25	0.25	0.25	0.25	0.25
20		0.25				0.50	0.13						0.50	0.08		0.50				0.50	0.17						0.50	0.08		0.50		0.25	0.50	0.50	0.25						0.50	0.08
23							0.00					0.25		0.04				0.25		0.25	0.08						0.30	0.15					1.00		0.17							0.00
24	0.25	0.25	0.75	0.25		0.75	0.38	0.25		0.25			0.75	0.21	0.25	0.25	0.25	0.75	1.00	0.75	0.54	0.25	0.25	0.25	0.25		0.75	0.29	0.25	0.25	0.25	0.25	0.50	0.75	0.38	0.25	0.25	0.25	0.25		0.75	0.29
25	0.25	0.25	0.25	0.25	0.25	0.50	0.29	0.25		0.25	0.50	0.25	0.50	0.29	0.25	0.25	0.50	0.50	0.25	0.50	0.38	0.25	0.25	0.25	0.25	0.25	0.40	0.28	0.25	0.25	0.25	0.25	0.25	0.50	0.29	0.25	0.25	0.25	0.25	0.25	0.50	0.29
26							0.00		0.25					0.04	0.25		0.50	0.50	0.25		0.25						0.30	0.15		0.25					0.04							0.00
27	0.25			0.25	0.50	0.50	0.25	0.25	0.25	0.25	0.50	0.25	0.50	0.33	0.50	0.50		0.75	0.50	0.50	0.46				0.75	0.50	0.50	0.29	0.25		0.25		0.50	0.17				0.75	0.50	0.50	0.29
28		0.25		1.00	1.00	0.50	0.46				1.00	0.75	0.50	0.38		0.50		1.00	1.00	0.50	0.50				1.00	1.00	0.50	0.42		0.50		1.00	1.00	0.50	0.50						0.50	0.08
31	0.25	0.25	0.25	0.25	0.50	0.50	0.33	0.25	0.25	0.50	0.25	0.50	0.50	0.38	0.25	0.25	0.25	0.25	0.50	0.75	0.38	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.25	0.25	0.25	0.25	0.50	0.29	0.25	0.25	0.25	0.25	0.25	0.50	0.29
32	0.25			0.25	0.25		0.13		0.25		0.25	0.25		0.13	0.25	0.25	0.25	0.25	0.50	0.25	0.29	0.25	0.25	0.25	0.75	0.50	0.65	0.44	0.25	0.25	0.25	0.75	0.50	0.25	0.38	0.25	0.25	0.25	0.25	0.25	0.25	0.25
33				0.50	0.25	0.75	0.25	0.25	0.25	0.25	0.25	0.25	0.75	0.33		0.50		0.50	0.50	1.00	0.42			0.50	1.00	1.00	1.00	0.58	0.25	0.25	0.25	0.75	0.75	0.75	0.50						1.00	0.17
34					1.00	1.00	0.33		0.25	0.50	0.50	1.00	1.00	0.54		0.50	0.75	1.00	1.00	1.00	0.71		1.00	1.00	1.00	1.00	0.67					1.00	1.00	1.00	0.50			0.75	1.00	1.00	1.00	0.63
35	0.25	0.25	0.75	0.25	0.75		0.38	0.25	0.25		0.50	0.75		0.29	0.25	0.25	0.50	0.50			0.25	0.25	0.25	0.25	0.75	0.75		0.38	0.25	0.25	0.25	0.75	0.75		0.38	0.25	0.25	0.25	0.75	0.75		0.38
36		0.25					0.04							0.00		0.50					0.08				0.50	0.25	0.30	0.28							0.00							0.00
38	0.25		0.25	0.50	0.50	0.25	0.29	0.25	0.25	0.25	0.50	0.50	0.25	0.33	0.25		0.25	0.50	0.50	0.75	0.38	0.25	0.25	0.25	0.50	0.50	0.75	0.42	0.25	0.25	0.50	0.50	0.50	0.25	0.38	0.25	0.25	0.50	0.50	0.25	0.25	0.33
39	0.50	0.75	0.25	0.50	0.50	0.75	0.54	0.50	0.50		0.50	0.50	0.75	0.46	0.50	0.25	0.25	0.50		0.75	0.38	0.50	0.75	0.25	0.50	0.50	0.75	0.54	0.50	0.25	0.25	0.50	0.50	0.75	0.46	0.50	0.75	0.25	0.50	0.50	0.75	0.54
40	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.25	0.25	0.25	0.25	0.75	0.33	0.25	0.25	0.75	0.75	0.75	0.75	0.50	0.25	0.25	0.25	0.75	0.25	0.75	0.42	0.25	0.25	0.25	0.25	0.25	0.75	0.33
Avgl	0.14	0.13	0.18	0.22	0.29	0.38	0.22	0.16	0.15	0.15	0.25	0.26	0.40	0.23	0.17	0.26	0.28	0.41	0.39	0.44	0.33	0.14	0.15	0.16	0.42	0.39	0.69	0.32	0.15	0.25	0.15	0.38	0.37	0.46	0.28	0.13	0.13	0.15	0.22	0.22	0.49	0.22

6.3. Teaching and Lecture Notes to Introduce Strategies and Tools to Evaluate Rubrics

The assertions map depicted in Table 3.2 is useful for the CAD trainer in order to prepare a sequence of tasks aimed at maximizing quality comprehension, but it is hypothesized that it is also important to instruct the CAD trainees in quality concepts from the beginning of their instruction. Hence, the assertions map and the derived rubrics were initially made publicly available for the CAD trainees from the beginning of their training. Therefore, they knew that comprehension of those quality criteria was mandatory in order to pass the exams. However, as a result of previous tests, it was realized that making this rubric information available was not sufficient for the students to understand and utilize the concepts embedded in these rubrics. The next step was explicitly explaining the meaning of every rubric before requiring the students to use them.

The third experiment was aimed at determining which specific teaching and lecture notes, if any, are helpful for the students to understand rubric assertions. It was hypothesized that both suitable strategies and evaluation tools had to be conveyed. Currently, the author only considers tools already built in CAD applications, and uses them to develop metrics aimed at helping to understand and mark those rubric assertions that are difficult to quantify and evaluate. It is suggested that the students use them primarily as assessment tools (regardless of its intended main use).

To this end, suitable explanations for each assertion was elaborated. For instance, to illustrate assertion M3.2 (profiles flexibility) it is argued that profiles are flexible if: 1) profile constraints are weak enough to allow local changes, and 2) profile constraints are strong enough to prevent local changes from causing undesired changes or errors. But, both goals are contradictory, so an agreement is

required. For instance, in order to convert a quadrilateral into a square, different strategies are equally valid: 1) make two opposite lines horizontal and the other two vertical, and make two consecutive lines equal (Figure 3.7 left), or 2) make three out of the four consecutive lines perpendicular, and make two consecutive lines equal (Figure 3.7 right). Both sets constrain the shape, but not the position. However, it is important to note that the two sets of constraints are not equivalent: the first set adds an extra constraint between the shape and the reference system and prevents rotation, while the second set allows rotation. The difference between weak and strong constraint sets can be further emphasized by noting that making opposite sides parallel to each other for the first approach is: 1) insufficient on its own (too weak), or 2) redundant, if previous constraints are already present (too strong).

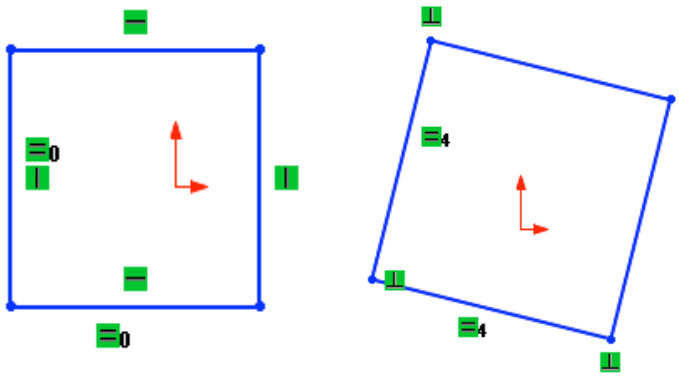


Figure 3.7. Profiles with different constraints result in different flexibility.

To summarize, profiles are flexible if shape, size, position and orientation are fixed independently (Figure 3.8).

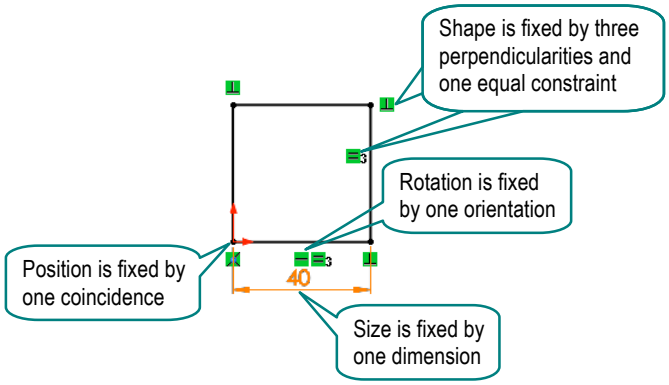


Figure 3.8. Profiles are flexible if shape, size, position and orientation are fixed independently.

As can be deduced from the above example, current teaching and lecture notes on rubrics are a combination of check-lists, good practices, and evaluation tools, all of them related through assertions and sequenced according to the assertions map. However, it can be noted that differences that exist with previous approaches are important. Instead of replacing best-practices manuals with checklists, as advocated by Gebhard [2013], best practices are used to explain why items in the checklist are important. Instead of introducing best-practices and/or checklists after the training period, they are inserted along the training period by way of rubrics.

The most important difference is that checklists are only valid for the most basic dimensions of quality: validity and completeness. The other dimensions always imply trade-offs between opposite criteria, such as the quite obvious incongruity between robustness and flexibility in regards to consistency, or the less obvious inconsistency between conciseness and simplicity. These types of trade-offs are at the core of the design process so students should be taught about them as soon as possible and they should be given clear criteria in order to find reasonable solutions. For instance, as said above, a profile is robust if it is fully constrained and is flexible if shape, size, position and orientation are fixed independently. The profile is consistent if both properties are simultaneously achieved.

Finally, the author advocates the use of tools already available in most CAD applications as evaluation tools. A trivial example would be "pushing" lines or vertices to try to distort an unconstrained square (by way of editing tools) as a simple method to discover which constraints are missing (Figure 3.9). This approach is much more intuitive and direct than trying to count degrees of freedom. However, it is also noted that a new tool aimed at highlighting unconstrained degrees of freedom or unnecessary mutual dependencies could be beneficial for novice users. So, there is room for new tools aimed at training novice product designers in quality-oriented strategies in CAD.

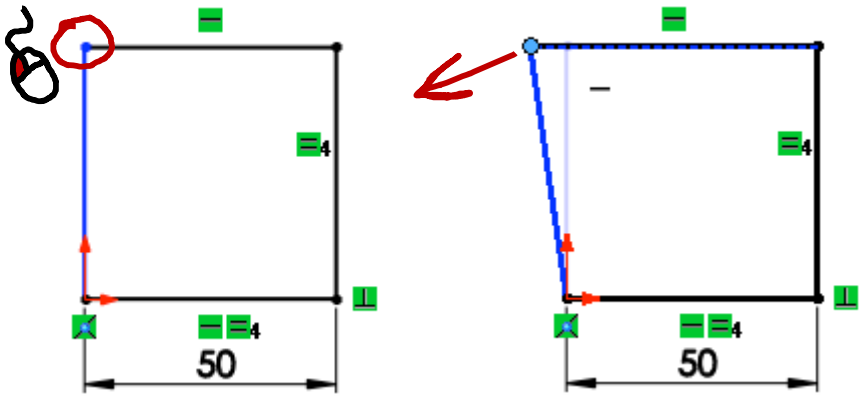


Figure 3.9. Pick and try to drag different lines and nodes to visualize unconstrained degrees of freedom.

According to the pilot experiment, students exposed to such explanations understand the quality concepts embedded in the rubrics, especially as they obtain self-evaluations which are beneficial and much more similar to teacher evaluations than non-exposed students (see Table 3.7). The experiment is described as follows: after being taught about the meaning of the first rubric in the assertions map, and after a two-hour session in the CAD laboratory where they solved very simple profiles (such as a quadrilateral given the lengths of four sides and one angle), the students were required to solve the profile displayed in Figure 3.10, where dimension L (73.30) had to be obtained as a derived value and was used by the students to check the integrity of their profiles. The exercise solution was provided to all students, but they were discouraged from consulting it unless it was needed. After completing the exercise, they were asked to self-evaluate their job using the rubric displayed in Figure 3.11. Then, the teacher evaluated their performance using identical rubrics and both evaluations were compared.

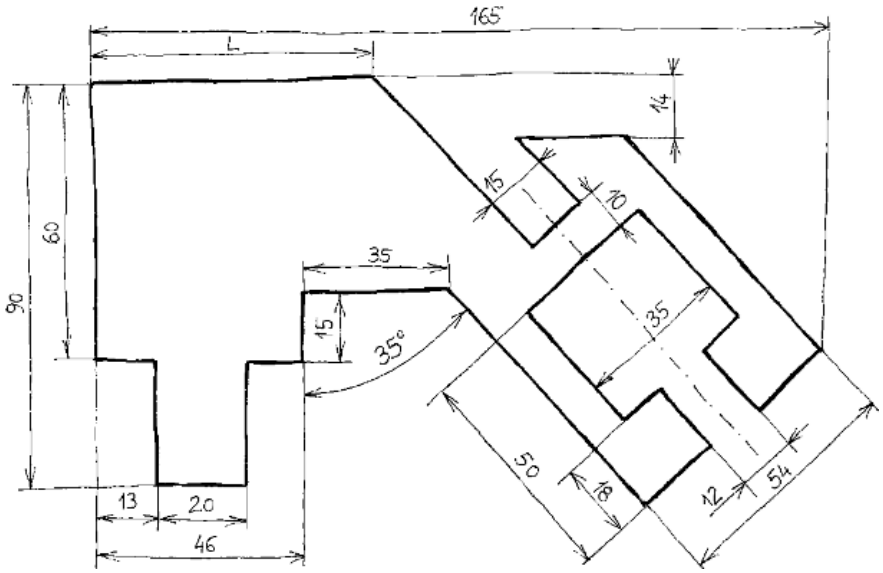


Figure 3.10. Complex profile to draw and self-evaluate with rubrics during Practice 1.

Student:_____

For each row, mark the cell which is closest to your opinion about the assertion

Item	Assertion	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1	**The model is valid**					
1.1	Model can be retrieved					
1.1a	Model is not missing (neither is lost nor wrongly named)					
1.1b	There are no errors opening document of the model					
1.2	Model can be used					
1.2a	Models do not contain error messages					
1.2b	Document of the model is compatible with the application					
3	**The model is consistent**					
3.1	Profiles are robust (changes do not produce unexpected failures)					
3.1a	No one profile does contain duplicated lines					
3.1b	No one profile does contain segmented lines					
3.1c	Profiles are always fully constrained					
3.2	Profiles are flexible (allow many changes)					
3.2a	Profile constraints are weak enough to allow local changes (design variations)					
3.2b	Profile constraints are strong enough to prevent local changes (design variations), from causing undesired changes or errors					

Figure 3.11. Scoring Rubric of Practice 1.

20 students solved the exercise and 18 returned the form, with two forms being incomplete and one student solving the exercise with a "future version" of the CAD application. These five were discarded and the remaining 15 valid self- evaluations were compared against the teacher evaluations (Table 3.8).

The primary observed result is that the students clearly understand the rubric, as differences between teacher evaluation and self-evaluation are negligible. Absolute values of the teacher evaluations are

also relevant, as they illustrate that students obtained good marks because they had clear understanding of what validity and consistency of profiles mean (Table 3.9).

Differences between self-evaluation and teacher evaluation reveal two important issues. First, differences in Assertion 1.1b were due to students not realizing that an open sketch is vulnerable: it may be inadvertently modified while re-opening the document. As a result of the pilot experiment, suitable explicit explanations on the importance of exiting profiles were added to the corresponding lecture notes. Second, although the differences between teacher and student evaluations were small, and the teacher evaluation was good in general, the trade-off between Assertions 3.2a and 3.2b was not fully understood by all students. More time is required to resolve such trade-offs.

Finally, it is noted that Subject 19 was assessed by the teacher with the highest mark in all the assertions. Hence, the discrepancy in Assertion 1.1a seems to be due to misunderstanding the assertion, as he replied as "strongly disagree," while he delivered a perfectly valid model in a valid file. Possibly, the implicit double negative in the sentence caused confusion. Attempts should be made to prevent such complex assertions in future rubrics.

Table 3.8. Differences between the self-evaluation and the teacher evaluation in Experiment 3 (void means no difference; 1 means 100% difference).

Student	1.1	1.1a	1.1b	1.2	1.2a	1.2b	3.1	3.1a	3.1b	3.1c	3.2	3.2a	3.2b	
1			0.25								0.25		0.25	
2												0.25		
3											0.25	0.25		
4			0.25								0.5		0.25	
8			0.25								0.25		0.25	
9			0.25				0.25	0.25	0.25	2	2		0.25	0.40
10											0.5	0.25	0.5	
11			0.25											
12												0.25		
15			0.25									0.25		
16												0.25		
17											1	0.25		
18			0.25								0.25	0.25		
19		1												
20											0.75	0.75		
Avg|											0.38			0.08

Table 3.9. Teacher evaluation in Experiment 3 (Strongly agree=1, Strongly disagree=0).

Student	1.1	1.1a	1.1b	1.2	1.2a	1.2b	3.1	3.1a	3.1b	3.1c	3.2	3.2a	3.2b
1	1	1	0.75	1	1	1	1	1	1	1	1	0.75	1
2	1	1	1	1	1	1	1	1	1	1	1	0.75	1
3	1	1	1	1	1	1	1	1	1	1	1	1	0.75
4	1	1	0.75	1	1	1	1	1	1	1	1	0.5	1
8	1	1	0.75	1	1	1	1	1	1	1	1	0.75	1
9	1	1	0.75	1	1	1	0.5	1	1	0.5	0.5	0.75	0.5
10	1	1	1	1	1	1	1	1	1	1	1	0.75	1
11	1	1	0.75	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	0.75	1
15	1	1	0.75	1	1	1	1	1	1	1	1	0.75	1
16	1	1	1	1	1	1	1	1	1	1	1	0.75	1
17	1	1	1	1	1	1	1	1	1	1	1	0.75	1
18	1	1	0.75	1	1	1	1	1	1	1	0.75	0.75	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	0.75	0.75	1

Results of the pilot experiment suggest that current teaching and its associated lecture notes on rubrics are working reasonably well and the intention is to develop them in detail. They include checklists, as much as explanations on typical mistakes, but they are centered around good practices and evaluation tools, all of them related around rubric-assertions and sequenced according to the assertions map. The most important innovations are: the main role played by explanations of how to cope with trade-offs between contradictory criteria and the explanation of how existing tools may be used to assist in evaluating assertions.

7. Conclusions

The author intends to convey quality-oriented strategies to CAD trainees, so quality criteria are embedded into the rubrics in order to force CAD trainees to comprehend them early in their instruction and prevent the Einstellung effect.

But producing and using such rubrics is not trivial. The author was a member of a research team aimed at providing measuring and validation tools to improve rubrics for quality-oriented CAD training. The team first identified those aspects of quality in CAD models that are appropriate to introduce to large populations of inexperienced CAD trainees, and grouped them around six main dimensions (Table 3.1). Then, an expand-contract strategy that adapts the rubrics to CAD trainee progress was introduced (Table 3.2). As a result, an assertions map was obtained. This map shows how the expand-contract strategy adapts the rubrics to the progress of the CAD trainees and helps them to understand the different dimensions of the rubrics. In this chapter, the author's improved version of this approach has been described, while the main contribution is an improved arrangement and definition of the assertions for modeling described in Table 3.1.

This chapter also included several pilot experiments aimed at illustrating the suggested process for replicating an approach for further developing rubrics adapted to other scenarios. The results of these pilot experiments were useful to a) discourage the use of interspersed rubrics and suggest that each task must be evaluated by a separate rubric (although the results were not fully conclusive); b) suggest

that the expand-contract process is beneficial to comprehend the quality concepts embedded in the rubrics (i.e. learning to work with generic rubrics is not enough, nor is simply being exposed to quality rubrics); and c) realize that the students do not understand and utilize the concepts embedded in the rubrics unless the meaning of every assertion is explicitly explained and they are forced to use them and compare their self-evaluations against the instructor evaluations. It was also determined that while some "simple" assertions are well understood by most of the students after initial exposure, others still require further training in order for them to be comprehended. Hence, the expand-contract process should be evaluated in a full-scale experiment aimed at finding the right contraction speed for each assertion. It is also realized that students require explicit procedures and metrics to assist them in evaluating their performance.

The last experiment was aimed at determining whether explicit explanations are helpful in conveying suitable CAD modeling strategies. The results suggest using evaluation tools built in CAD applications to develop metrics aimed at helping to understand and mark those assertions of rubrics that are difficult to quantify and evaluate. Current teaching and lecture notes on rubrics appear to work reasonably well and should be developed in detail. They are a mix of checklists, good practices, and evaluation tools, all of them related to assertions and sequenced according to the assertions map. They pay particular attention to the trade-offs required to obtain valid designs.

Chapter 4
The Implementation of Part Rubrics

1. Introduction

Effective learning of Mechanical Computer-Aided Design (MCAD) still remains a major challenge today in both academic and industrial settings. Current educational practice is primarily oriented to provide students with declarative and specific procedural command knowledge as classified by Chester [2008]. Declarative command knowledge is related to generic commands or algorithms that usually are available within MCAD systems, such as geometric transformations (i.e. patterns and mirroring operations), or solid modeling operations (i.e. extrusions or revolutions). The practical application of this knowledge to a particular CAD system requires specific procedural command knowledge, which is system and version specific. Usually, procedural command knowledge centers the teaching/learning activities in the typical MCAD courses. However, this approach provides neither enough strategic knowledge, which is associated to the election of the proper modeling schemes [Toto et al., 2014], nor is explicitly focused on quality.

The concept of quality in the realm of MCAD is a relatively recent topic of interest [Contero et al., 2002], [Gu et al., 2001], and [Yang et al., 2006]. Its more developed aspect relates to the geometrical and topological correctness of models, that can be evaluated with ad hoc software (quality checkers) that implement quality criteria as defined in standards such as VDA 4955 [1999] or SASIG PQD 2.1 [2005]. However, higher level dimensions of quality, as those related to CAD model reusability [Otto and Mandorli, 2015] and [Mandorli et al., 2015] that are identified as semantic quality in [Contero et al., 2002] rarely are covered in regular CAD courses. This level of quality is intimately connected with the effective application of strategic knowledge. In industrial settings, this process is commonly founded on the identification and application of best-practices, which are circulated electronically within the company. Unfortunately, most best-practices documents are developed in-house and are not released publicly. Similarly, modeling methodologies [Camba et al., 2014] that define systematic ways for the construction of the CAD models are also commercially unavailable.

Historically, a series of quality dimensions have been defined [Company et al., 2015] and they provide a framework where a set of attributes to measure the different components of quality are defined. The evaluation of the quality of CAD models against a set of quality criteria with different levels of performance can be naturally adapted to the rubric concept employed in educational settings.

Rubrics can be used to support an objective scoring process (supplying consistent grading criteria), which can then be shared between instructors involved in a common subject [Mertler, 2001]. Rubrics can also be applied to support self-evaluation and peer review [Andrade and Du, 2005] and [Karkehabadi, 2013]. In this context, rubrics provide a transparent framework to publicize expectations of quality on assessed tasks, but many authors argue that rubrics can teach as well as evaluate [Andrade and Du, 2005] and [Bailey and Szabo, 2006]. In fact, the formative use of scoring rubrics has been proven useful under some assumptions [Panadero and Jonsson, 2013].

Computer Aided Design (CAD) systems are complex software tools whose learning requires correctly managing the functionality provided by the system while applying proper strategies to create high quality CAD models ("content-based constructed responses" [Liu et al., 2014]). CAD models are a relevant part of industrial product design, which is a complex process that requires evaluation [Bailey and Szabo, 2006] and [Waks et al., 2011]. Formative rubrics convey quality-oriented strategies to CAD trainees since quality criteria can be embedded in scoring rubrics to force them to grasp these concepts early in their instruction [Company et al., 2015]. This strategy prevents the natural tendency to solve problems in a particular way, ignoring better alternative methods (Einstellung effect), which has been previously described in the particular case of CAD training [Mandorli and Otto, 2013].

Nevertheless, simply using paper-based formative rubrics for CAD model quality is not enough. Computer-assisted grading methods are essential to improve the efficiency and effectiveness of assessment through rubrics [Anglin et al., 2008] and [Auvinen et al., 2009]. The use of static rubrics are not appropriate in CAD, as common problems do not lend themselves to only "right" or "wrong" answers [Reid and Cooney, 2008]. CAD models may be seen as open-ended problems and they require sophisticated evaluation strategies [Diefes-Dux et al., 2010]. Computer Assisted Assessment (CAA) provides automatically contextualized feedback [Santos et al., 2009] and [Cebrián-Robles et al., 2014] and some CAA tools are embedded into proprietary learning management systems (LMS) [Atkinson and Lim, 2013] and [Isbell and Goomas, 2014], which may prevent adoption as they may be complicated, costly, and non-customizable. It has also been argued that as students learn at a different pace, rubrics should be adaptable on demand for each student [Company et al., 2015].

Thus, the team's vision is that those requirements are primarily served by adaptive and adaptable rubrics, which are complementary concepts also used in other contexts (see, for instance [Miller et al., 2005]). While an adaptable rubric can be modified by the user to adapt to different needs, an adaptive rubric is able to adapt or change itself, depending on the pattern of use it undergoes. The origins of adaptive rubrics are briefly summarized in [Georgiadou et al., 2006], which also describes the main quality parameters for a good Computer Adaptive Testing (CAT) system. Adaptive rubrics have not yet been considered for CAD quality, although some studies are certainly worth citing. It is commonly known that instructors can customize adaptive rubrics according to the evolution of students' learning [Isbel and Goomas, 2014], [Rejab et al., 2011], and [Silva and Restivo, 2012]. Some systems can automatically score each criterion by collecting related data from the files of the CAA tools [Dimopoulos et al., 2013] and [Rayón et al., 2014], or from previous evaluations [Matthews et al., 2012] and [Cabrera and Villalon, 2013]. Other approaches allow users to customize the assessment process to a particular environment (allotted time, interacting device, etc.) [Harchay et al., 2014]. Finally, spreadsheets have been used to create adaptive rubrics [Nash, 2015]. Following Economides et al, the majority of the CAT systems "do not offer to the examinee any advanced support and functionalities" [2007]. In particular, they state that "the presentation and adaptation could be improved by personalizing the test to the examinee's personal taste". Hence, it can be concluded that adaptable rubrics, which could be modified by the user to adapt to different needs, would be clearly helpful.

There are two types of dynamic rubrics, which are developed (or re-defined) in this current research: while an adaptive rubric is instructor-driven (configured to automatically change itself depending on the usage pattern), an adaptable rubric is user-driven (interactively modified by the user to adapt to different needs). However, most currently available computer-assisted rubrics only deal with static ones [Malini and Andrade, 2010]. Hence, a new computer-assisted rubric aimed at conveying quality

concepts to CAD trainees has been developed. The new rubric provides feedback (shows comprehensive scores if requested), and is adaptable to the learning pace of each student, as different levels of detail can be deployed on demand according to student needs. Thus, if one criteria in the rubric is overly abstract or difficult to understand, additional levels of detail can provide a more easily understandable description of expected performance levels.

The design specifications, strategy implementation, and the lessons learned while developing and testing the prototype adaptable rubric are first described in this chapter. Preliminary conclusions derived from the pilot experiments through the prototype rubric are also described, then the process of developing electronic forms (e-rubrics) allowing for easy-to-create (computer-assisted) and easy-to-use (adaptable) formative rubrics are finally discussed and tested.

2. Related Work

Previous research on the use of rubrics to capture judgement is wide-ranging and general concepts on the topic are thoroughly discussed by Mertler [2001] and Karkehabadi [2013]. A majority of studies report that rubrics in higher education are used only for evaluation [Malini and Andrade, 2010], but they can also be utilized for acquiring or reinforcing complex skills [Manson and Olsen, 2010] and [Smit and Birri, 2014], or for assessment of engineering content-focused challenge questions [McCormick et al., 2015]. In other words, rubrics are suitable instruments for enhancing the psychometric properties of performance assessment, but also for supporting the process of formative assessment [Pophan, 1997], where assessment information is used to advise students about progress and aid in their development [Panadero and Jonsson, 2013].

Holistic rubrics score the overall process or product as a whole without judging the component parts separately [Nitko, 2001]. Hence, analytic scoring rubrics should be used for formative feedback [Mertler, 2001] as they allow scoring separate individual parts of the product or performance [Moskal, 2000].

In the specific case of formative analytic assessment in quality oriented CAD models, various rules and strategies to promote CAD model reusability have been described [Company et al., 2014] and the problem of improving communication of design intent has also received some attention [Otey et al., 2014]. It has also been shown that rubrics are supportive tools to convey quality criteria from the onset of instruction of future mechanical CAD users [Company et al., 2013].

However, “pacing” is important, as rubrics should adapt to the learning rhythm of each student, and this matter continues to be open problem. Instructors can customize adaptive rubrics according to student learning evolution [Isbell and Goomas, 2014], [Rejab et al., 2011], [Silva and Restiva, 2012], and [Matuk et al., 2015]. There are also systems that automatically score each criterion by collecting related data from the files of the CAA tools [Dimopoulos et al., 2013] and [Rayón et al., 2014], or from previous evaluations [Matthews et al., 2012] and [Cabrera and Villalon, 2013]. Moreover, some authors describe approaches which allow the user to personalize the assessment process with a particular environment (time available, interacting device, etc.) [Harchay et al., 2014]. But of primary interest is adaptable rubrics, which should allow students to regulate their pace of learning.

In particular, the intent is to let users adjust the level of detail in the description of the evaluation criteria. Performance criteria descriptors are a critical component of rubric design that merit thorough consideration [Tierney and Simon, 2004]. Advanced trainees do not require detailed explanations of each evaluation criteria, while more detail may be desirable for novices. The target rubric should allow students to display more detailed levels in the evaluation criteria (with their corresponding sub-scoring) on demand, in order to provide supportive information for learning and performance. Consistency of self-assessment scoring should be guaranteed for each different combination of detailed-level responses. Nash [2015] used spreadsheets (Google Forms) to create an adaptive rubric. A spreadsheet rubric has also been tested in the pilot study, in order to develop a prototype of an adaptable rubric while also evaluating its possible benefits.

3. Specifications of an Adaptable Spreadsheet Rubric

The specifications for a prototype of an adaptable rubric were: (1) it must allow for varying levels of detail, and (2) it must provide immediate feedback of the evaluation scores. The first specification implies that every student utilizing the rubric to score their CAD model should be able to select the level of detail that is required. Levels of detail, as described in [Company et al., 2015], where the so-called assertions map distinguishes between main criteria (Level 1) and two nested levels of sub-criteria (Level 2 and Level 3) are used. Level 1 is the highest level of abstraction, as it directly evaluates the six main dimensions associated with the quality of CAD modelling defined by [Company et al., 2015]: valid, complete, consistent, concise, easy to use, and conveyance of design intent (although "easy to use" has been reworded as "clear"). Levels 2 and 3 progressively give more details for a particular dimension.

The second specification derives from former experience with rubrics, where it was observed that some students tend to mark rubrics incoherently when working at different levels. Hence, it is proposed that scores must be automatically updated by the spreadsheet-rubric every time the user marks some criterion.

Three other design specifications were also considered: (3) the form should be easy to reconfigure and adapt, so as to allow the instructor to adapt the rubric to changing scenarios; (4) the information of the filled forms should be easily extracted and processed, and (5) the form should prevent incomplete or inconsistent scoring.

4. Prototype of Adaptable Spreadsheet Rubrics

To facilitate the requirements while producing a fast prototype, a spreadsheet configured to hide internal calculations, while simultaneously showing the list of criteria and the cells to score them through combo boxes was used. Active X controls and Visual Basic for Applications (VBA) were used. A short satisfaction test was included after the rubric. The result is shown in Figure 4.1.

As a result of the first two experiments (see next section), the spreadsheet was changed, replacing the combo boxes by radio buttons, hiding the score (which could now be shown on demand), and removing the satisfaction test (Figure 4.2). Radio buttons were discovered to be better than combo boxes because students declared that activating the drop-down list is a time-consuming task (some students required guidance to understand how the prototype rubric works) and it was realized that radio buttons can (while textboxes cannot) be visually inspected and compared after re-opening the document.

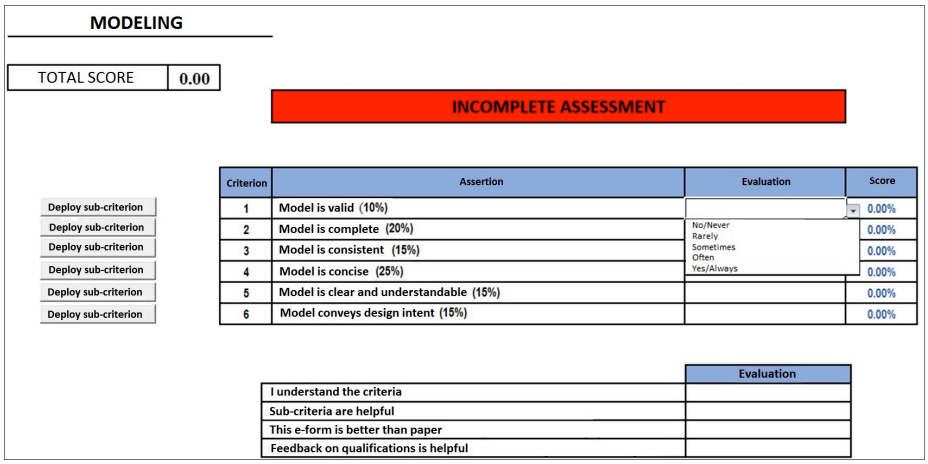


Figure 4.1. Front end of the prototype adaptable rubric showing the combo box for Criterion 1.

In both cases, and to allow for personalized levels of detail, the six criteria of Level 1 are displayed by default. Simply pushing the corresponding button, students are allowed to deploy Level 2 criteria for every particular main dimension. In doing so, they can score the Level 2 criteria or repeat the query for more detail, so as to get Level 3 criteria.

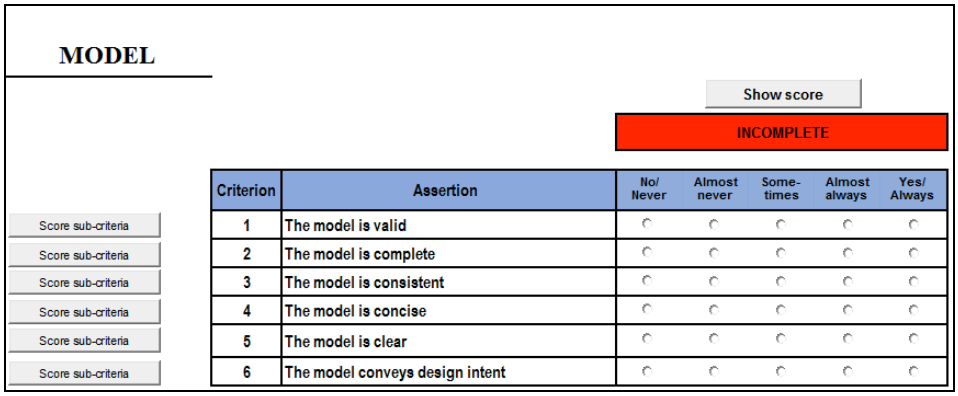


Figure 4.2. Front end of the prototype adaptable rubric with radio buttons.

As illustrated in Figure 4.3, once Level 2 criteria for Dimension 1 are shown, Criterion 1 is blocked and the student can only mark Sub-criteria 1.1 and 1.2. In doing so, the score for Criterion 1 is automatically calculated and displayed, as is the total score.

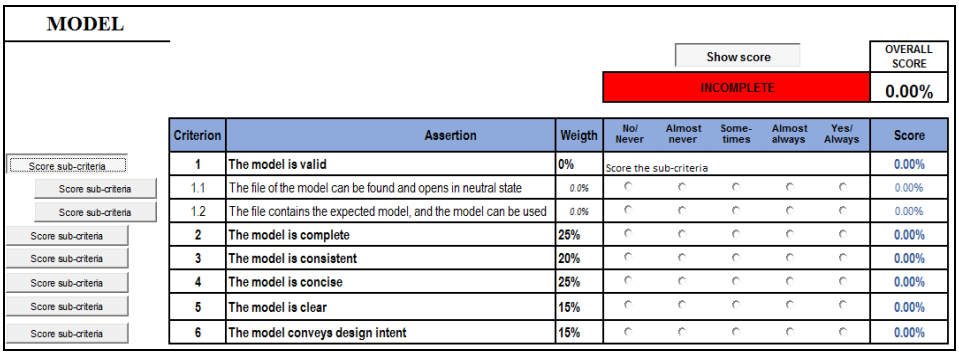


Figure 4.3. Illustration of how Level 2 is shown for Level 1 first criterion.

In order to extract information from the rubrics, an application aimed at sequentially opening all spreadsheets into a particular folder, read their answers, and save them into a single spreadsheet was developed (Tables 4.1, 4.2, and 4.3).

5. Experiments

Three experiments were conducted. The first demonstrated that results using paper-forms and spreadsheet-forms were similar, the second demonstrated that students using spreadsheet-forms felt free to use different levels of criteria when they need them, and the third demonstrated that when students were allowed to freely select between paper-forms and spreadsheet-forms, they greatly opted for the latter. Detailed descriptions on each experiment follow.

5.1 Experiment 1

Third year students of a four year mechanical engineering program were exposed to the prototype rubrics. They had been progressively introduced to the rubrics since the beginning of the term. The introduction included explanations about the meaning of the six quality dimensions (Level 1) and subsequent explanations on the detailed criteria to measure the degree of accomplishment of such dimensions (Level 2 and Level 3).

The students had been encouraged to self-mark their homework using paper form rubrics, but no formal evaluation was conducted, nor was their self-evaluation supervised by the instructors. The task was to model a single part of a simple assembly. Two sub-groups were created. Group 1 was required to model the ring depicted on the left side assembly of Figure 4.4. Groups 2 and 3 (which had different instructors) had to model the base plate depicted on the right side assembly of Figure 4.4. In both cases, the remaining parts were available so that students could freely dimension their models, but guaranteeing compatibility with the other parts.

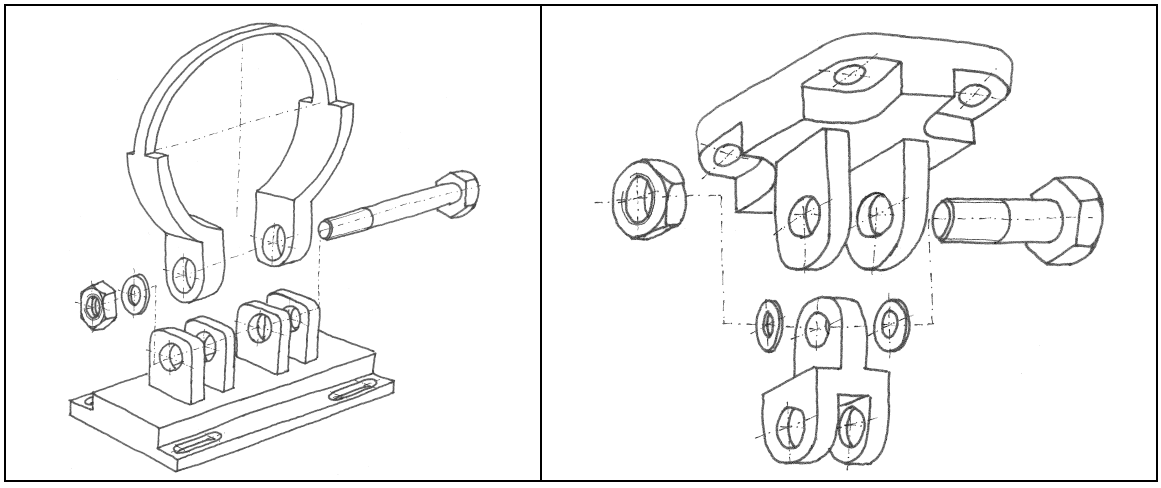


Figure 4.4. Assemblies used for modeling in Experiment 1.

Prior to the exam, all students were informed about the exact rubric that was to be used to evaluate their tasks. During the exam, they were asked to self-evaluate their own tasks using paper-forms. A few days later, they were allowed to voluntarily repeat their self-evaluation using the prototype electronic form. They were informed that only their best self-evaluation (the self-evaluation closest to the instructor evaluation) would be used for marking. At the same time, students were asked to peer-evaluate the work of another anonymous student, selected randomly by the instructor. They were informed that their evaluation would have no effect on the final mark of the evaluated student, but

would linearly increase their own marks from zero (if their peer-evaluation diverged from the instructor evaluation by more than 50%), up to 10% of the highest score (if their peer-evaluation was exactly equal to the instructor evaluation).

Forty-seven students were examined, but two of them did not return valid exams (Students 3 and 45). The remaining forty-five returned the paper rubrics, but only thirty of them also returned spreadsheet-form rubrics. All thirty students were analyzed together, as differences between sub-groups were found to be irrelevant. In comparing the self-evaluations of those thirty students, it was discovered that spreadsheet-forms were closer to the instructor evaluation for sixteen of them, while paper rubrics were closer for the other fourteen students. Differences between paper and spreadsheet-form evaluations were less than 5%: average 0.48 and median 0.47. Hence, this pilot study concluded that spreadsheet-forms do not negatively affect the rubrics-based evaluation aimed at conveying CAD model quality.

5.2 Other results for Experiment 1

Detailed results for this experiment are shown in Table 1, which summarizes and compares the following information: evaluation of the instructor; student self-evaluation (both with paper-form and spreadsheet-form); peer evaluation (where every student anonymously evaluated another randomly selected student); and differences between self and instructor evaluation and between peer and instructor evaluation. All evaluation values are in the range [0, 10].

Students 4 and 44 were excluded, whose self-evaluations differ more than 25% from each other (grayed cells in column “Difference paper-e-form”). Similarly, Student 2 was excluded, whose evaluations (self and peer) differ both by more than 25% from instructor evaluations (grayed cell in column “Big differences in self and peer”). It is noted that negative values in the columns “Difference Self-Instructor” and “Difference Peer-Instructor” indicate students that self-scored worse than they should (“pessimistic”), while positive values correspond to optimistic subjects.

It appears that during the second self-evaluation (with spreadsheet forms), students used levels as required, while during the first evaluation (with paper), there was a bias toward higher levels. In comparing the detailed answers from all the rubrics (summarized on the right side of each sub-table) it is observed that for the forty-five paper self-evaluations, students used higher level criteria (thirty Level 1, four Levels 1/2, one Levels 2/3 and eight Levels 1/2/3). On the contrary, the thirty spreadsheet forms for self-evaluation included ten Level 1, six Level 2, four Level 3, four Levels 1/2, two Levels 2/3 and four all three levels. A similar distribution occurred for peer evaluation through spreadsheet form: eleven Level 1, six Level 2, four Level 3, six Levels 1/2, two Levels 2/3, and one all three levels. It is surmised that this bias toward more abstraction and less mixture of levels while using paper forms may be due to the rigidity of the paper. However, this experiment does not provide conclusive evidence as to whether students prefer paper or spreadsheet forms, since spreadsheet forms were only available for the second evaluation, where students had as much time as needed in order to review their task and complete the spreadsheet form. On the contrary, paper rubrics had been completed under pressure (at the end of the exam), and with less experience. So, it should be confirmed whether spreadsheet forms negatively affect the evaluation by using rubrics, in the absence of such circumstances.

Other conclusions may be derived from Table 4.1, where differences between self and instructor evaluations are 0.79 in average and 0.41 in median, and differences between peer and instructor evaluations are 1.12 in average and 0.62 in median. So, it can be easily concluded that evaluations of students align moderately well with evaluations of the instructor. In other words, differences may rise sporadically to more than 40%, but are typically smaller than 10%.

Table 4.1. Comparison between instructor, self, and peer evaluations for Experiment 1.

Group	Subject	Teacher	Self paper	Self e-form	Best self	Difference Self-Teacher	ABS difference Self-Teacher	Best form	Levels for self paper	Levels for self e-form	Peer	Peer e-form	Difference Peer-teacher	ABS difference Peer-Teacher	Levels for peer	Difference Paper - e-form	Big difference in self and peer	Subject to analyse	ABS Difference Self-Teacher	ABS Difference Peer-teacher
2	1	8.38	4.00		4.00	-4.38	4.38	Only paper	1		5									
1	2	9.09	6.13	5.50	6.13	-2.97	2.97	Paper	1/2/3	1	8	4.50	-3.97	3.97	1	0.63	2			
2	3																			
3	4	8.38	8.63	5.13	8.63	0.25	0.25	Paper	1/2/3	2	7	8.91	0.91	0.91	2	3.50				
3	5	6.47	7.38		7.38	0.91	0.91	Only paper	1/2/3		9									
1	6	3.50	3.13		3.13	-0.38	0.38	Only paper	1		10									
2	7	8.00	8.00	8.81	8.00	0.00	0.00	Paper	1	3	11	5.42	-0.58	0.58	3	-0.81		7	0.00	0.58
1	8	8.47	8.00	7.57	8.00	-0.47	0.47	Paper	1	1/2/3	15	9.53	-0.47	0.47	1/2	0.43		8	0.47	0.47
2	9	7.88	9.50	9.31	9.31	1.44	1.44	e-form	1/2	2	12	5.34	-0.66	0.66	1/2	0.19		9	1.44	0.66
1	10	7.25	9.00		9.00	1.75	1.75	Only paper	1		18									
2	11	6.00	5.75	7.63	5.75	-0.25	0.25	Paper	1	3	13	7.57	0.45	0.45	3	-1.88		11	0.25	0.45
2	12	6.00	7.00	5.38	5.38	-0.63	0.63	e-form	1	1	14	8.13	3.17	3.17	1	1.63		12	0.63	3.17
2	13	7.13	7.00	6.63	7.00	-0.13	0.13	Paper	1	1	16	7.63	-0.47	0.47	1	0.38		13	0.13	0.47
3	14	4.96	7.63		7.63	2.67	2.67	Only paper	1/2		17									
1	15	####	10.00	9.44	10.00	0.00	0.00	Paper	1/2/3	3	20	6.45	0.70	0.70	3	0.56		15	0.00	0.70
2	16	8.09	7.75	5.78	7.75	-0.34	0.34	Paper	1/2/3	2/3	19	7.47	2.66	2.66	2	1.97		16	0.34	2.66
3	17	2.94	3.50	3.97	3.50	0.56	0.56	Paper	1/2/3	2	23	8.70	-0.46	0.46	2/3	-0.47		17	0.56	0.46
1	18	7.06	9.25		9.25	2.19	2.19	Only paper	1		21									
2	19	4.81	10.00	9.28	9.28	4.47	4.47	e-form	1	1/2/3	24	9.22	2.13	2.13	1/2/3	0.72		19	4.47	2.13
1	20	5.75	6.88	7.03	6.88	1.13	1.13	Paper	1	1/2	22	9.03	-0.38	0.38	1/2	-0.16		20	1.13	0.38
1	21	7.59	8.50		8.50	0.91	0.91	Only paper	1		25									
1	22	9.41	6.50	6.84	6.84	-2.56	2.56	e-form	1	3	26	7.30	-1.45	1.45	3	-0.34		22	2.56	1.45
2	23	9.16	9.00	8.03	9.00	-0.16	0.16	Paper	1	2	27	5.63	-0.41	0.41	2	0.97		23	0.16	0.41
3	24	7.09	7.38	8.75	7.38	0.28	0.28	Paper	1	1/2	28	4.88	-2.13	2.13	1/2	-1.38		24	0.28	2.13
1	25	9.69	10.00	9.52	9.52	-0.17	0.17	e-form	1	1/2/3	29	8.88	1.09	1.09	1/2	0.48		25	0.17	1.09
1	26	8.75	9.63		9.63	0.88	0.88	Only paper	1		32									
2	27	6.03	8.88	9.13	8.88	2.84	2.84	Paper	1	1	30	5.13	0.00	0.00	1	-0.25		27	2.84	0.00
2	28	7.00	5.38		5.38	-1.63	1.63	Only paper	1		31									
1	29	7.78	6.25	7.55	7.55	-0.23	0.23	e-form	1	1/2/3	36	6.29	0.01	0.01	2/3	-1.30		29	0.23	0.01
3	30	5.13	6.00	5.88	5.88	0.75	0.75	e-form	1/2/3	1	33	6.75	-0.13	0.13	1	0.13		30	0.75	0.13
3	31	5.91	6.63		6.63	0.72	0.72	Only paper	1/2/3		35									
1	32	7.22	10.00		10.00	2.78	2.78	Only paper	1		37									
3	33	6.88	10.00		10.00	3.13	3.13	Only paper	1		38									
3	34	0.00	4.88		4.88	4.88	4.88	Only paper	1/2											
3	35	4.06	6.63		6.63	2.56	2.56	Only paper	1		39									
1	36	6.28	9.00	7.34	7.34	1.06	1.06	e-form	1	2	43	4.50	-2.56	2.56	2	1.66		36	1.06	2.56
1	37	8.16	10.00	8.75	8.75	0.59	0.59	e-form	1	1	47	4.13	0.44	0.44	1	1.25		37	0.59	0.44
3	38	6.91	4.88	6.63	6.63	-0.28	0.28	e-form	1/2	1	40	7.88	-0.34	0.34	1	-1.75		38	0.28	0.34
2	39	7.03									41									
3	40	8.22	9.25	7.38	7.38	-0.84	0.84	e-form	1	2/3	42	3.22	-1.97	1.97	2	1.88		40	0.84	1.97
3	41	7.41	7.75	7.38	7.38	-0.03	0.03	e-form	1	1	44	7.38	0.25	0.25	1	0.38		41	0.03	0.25
3	42	5.19		7.00	7.00	1.81	1.81	e-form		1	46	7.13	-1.03	1.03	1					
1	43	7.06	7.25	6.78	7.25	0.19	0.19	Paper	1	2	2	7.56	-1.53	1.53	2	0.47		43	0.19	1.53
2	44	7.13	8.88	6.28	6.28	-0.84	0.84	e-form	1	1/2	1	4.63	-3.75	3.75	1/2	2.59				
2	45																			
2	46	8.16	8.75	8.06	8.06	-0.09	0.09	e-form	2/3	1/2	4	6.75	-1.63	1.63	1	0.69		46	0.09	1.63
1	47	3.69	4.63	2.75	2.75	-0.94	0.94	e-form	1	1	16	5.13	-2.97	2.97	1	1.88		47	0.94	2.97
					Average		1.27	Paper 14				Average		1.29		0.48		Average	0.79	1.12
					Median		0.84	e-form 16				Median		0.80		0.47		Median	0.41	0.62

It can also be noted, that in general, students are optimistic about their own merits and pessimistic about the merits of the other students. When the student evaluation is higher than the instructor evaluation, their impression is referred to as "optimistic" and when the difference between the evaluations is negative, it is referred to as "pessimistic." Students are less optimistic the second time they self-evaluate their own task. Also, the students appear more accurately evaluate their own tasks

than the peers' tasks (average 0.79 and median 0.41 versus average 1.12 and median 0.62, respectively), conceivably because they are unable to guess how the model was built.

5.3 Experiment 2

In the second experiment, master's degree students were required to model one part depicted in a detailed drawing. They had been progressively introduced to rubrics from the beginning of the term. This introduction included explanations about the meaning of the six quality dimensions (Level 1), and subsequent explanations on the detailed criteria to measure the degree of accomplishment of such dimensions (Level 2 and Level 3). The part drawing, shown in Figure 4.5, was borrowed from Bertoline et al. [1999]. Self-evaluation was completed at the end of the exam, while peer evaluations were done a few days later. In both cases, master's students were provided with the prototype electronic forms. All students were informed that their final marks would increase linearly from zero (if their self-evaluation diverged from the instructor evaluation by 50% or more), up to 10% of the highest score (if their self-evaluation was exactly equal to the instructor evaluation). The same reward was applied for the peer-evaluation.

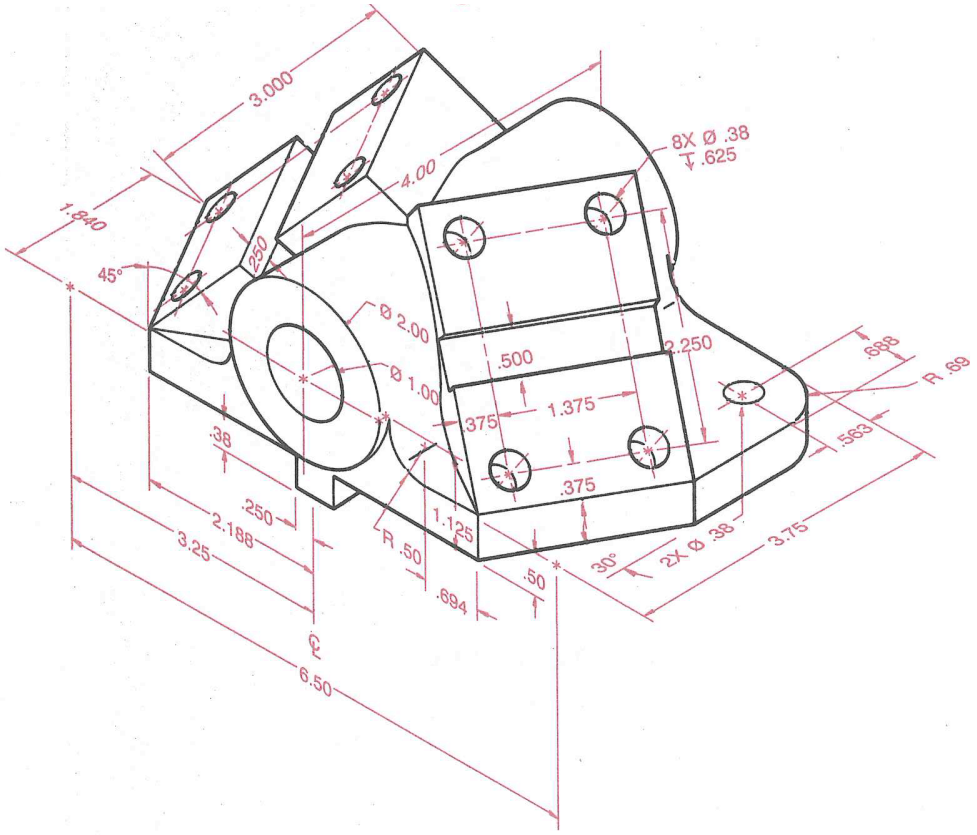


Figure 4.5. Detailed drawing of the Bertoline et al. part to model during Experiment 2.

All thirteen students returned valid self-evaluations, while only twelve returned valid peer-evaluations. Student 3 did not return a peer-evaluation, and differences between self and peer-evaluation were exceedingly high for Student 9, so they are not considered during the analysis. For the remaining students, it is noted that differences are higher for self-evaluation than for peer-evaluation. In the self-evaluation, up to six students differed more than 10% from the instructors, and four of them were more optimistic than the instructors. In the peer-evaluation, only one student differed more than 10%. Most likely, they were tired and short of time during the self-evaluation. During informal conversations (after the evaluations were collected), some students recognized that they had completed the self-evaluation from memory, without reviewing the actual contents of their own models. It is therefore concluded that students should only perform the evaluation well after the exam.

However, in comparing the detailed answers from all the rubrics (summarized on the right side columns in Table 4.2) it can be noted that only one master's student used Level 3 criteria to self-evaluate; five used Level 2 mixed with Level 1, and the other seven used only Level 1 criteria. There was no apparent correlation between the levels used and the accuracy of the evaluation. A similar behavior was noticed in the peer evaluations. Although a clear difference was perceived, students used, in general, additional detail for peer-evaluation than for self-evaluation. A matter of future interest could be examining if post-exam evaluation improves the efficiency of the evaluation. The main conclusion to be drawn is that spreadsheet forms allow the students to freely select the level they require.

5.4 Other results for Experiment 2

Detailed results for this experiment are shown in Table 4.2, which summarizes and compares the following information: instructor evaluation (two different instructors evaluated the models and average marks were calculated); student self-evaluation; peer-evaluation (where each student anonymously evaluated another randomly selected student), and differences between self and instructor evaluation and also between peer and instructor evaluation. All values are in the range [0, 1], except the values in bold, which are in the range [0, 10]. In calculating these totals, different weights were assigned to each dimension: 10% for Dimension 1, 20% for Dimension 2, 25% for Dimension 4 and 15% for the others. As shown in Table 4.2, all thirteen students returned valid self-evaluations, while only twelve returned valid peer-evaluations.

Other interesting results can also be obtained from Table 4.2. Maximum agreement is obtained for Dimension 1 (validity), since it is clearly the easier criterion to evaluate. On the contrary, maximum differences appear for Dimension 6 (conveys design intent), and Dimension 3 (consistency). Table 4.2 also shows that dispersion is minor (by average), but becomes important for some students and dimensions. So, it is concluded that additional work remains to be completed to obtain criteria that is easier to evaluate and more objective in nature.

Table 4.2. Comparison between instructor, self, and peer evaluations for Experiment 2.

Subject	1 (10%)	2 (20%)	3 (15%)	4 (25%)	5 (15%)	6 (15%)	**Average Profs.**	1 (10%)	2 (20%)	3 (15%)	4 (25%)	5 (15%)	6 (15%)	**Total Self**	Difference 1	Difference 2	Difference 3	Difference 4	Difference 5	Difference 6	**Dif. Self**
1	100	0.75	0.75	0.94	0.75	0.94	**8.50**	100	100	0.81	100	0.75	100	**9.34**	0	0.25	0.063	0.063	0	0.063	**0.84**
2	100	0.94	0.78	100	100	0.81	**9.27**	100	0.88	0.97	100	100	0.88	**9.52**	0	-0.06	0.188	0	0	0.063	**0.25**
3	0.88	0.25	0.63	0.25	0.50	0.13	**3.88**	0.00	0.00	0.50	0.25	0.50	0.50	**2.88**	-0.88	-0.25	-0.13	0	0	0.375	**-1.00**
4	100	0.94	0.83	0.69	0.44	0.25	**6.87**	0.75	0.75	0.75	0.75	100	0.75	**7.88**	-0.25	-0.19	-0.08	0.063	0.563	0.5	**1.01**
5	100	100	0.75	0.63	100	0.94	**8.59**	100	100	0.81	100	100	0.88	**9.53**	0	0	0.063	0.375	0	-0.06	**0.94**
6	100	0.31	0.81	0.75	0.63	0.56	**6.50**	0.75	0.63	0.25	0.75	0.50	0.50	**5.75**	-0.25	0.313	-0.56	0	-0.13	-0.06	**-0.75**
7	100	0.94	0.78	100	0.94	0.94	**9.36**	100	100	100	100	0.88	100	**9.81**	0	0.063	0.219	0	-0.06	0.063	**0.45**
8	100	0.94	0.56	0.75	0.75	0.63	**7.66**	100	100	100	0.75	0.50	0.75	**8.25**	0	0.063	0.438	0	-0.25	0.125	**0.59**
9	100	0.75	0.64	0.63	0.75	0.75	**7.27**	100	0.63	100	100	100	100	**9.25**	0	-0.13	0.359	0.375	0.25	0.25	**1.98**
10	100	0.63	0.63	0.75	0.88	0.69	**7.41**	0.75	0.50	0.50	0.75	0.75	0.50	**6.25**	-0.25	-0.13	-0.13	0	-0.13	-0.19	**-1.16**
11	100	0.94	0.53	0.75	0.88	0.88	**8.17**	100	0.75	0.75	0.50	0.75	0.75	**7.13**	0	-0.19	0.219	-0.25	-0.13	-0.13	**-1.05**
12	100	0.94	0.63	0.44	0.63	0.19	**6.13**	100	100	0.75	0.75	0.75	0.75	**8.25**	0	0.063	0.125	0.313	0.125	0.563	**2.13**
13	100	100	0.69	0.69	0.75	0.75	**8.00**	100	100	100	100	100	100	**10.00**	0	0	0.313	0.313	0.25	0.25	**2.00**
												Average			-0.1	0.0	0.1	0.1	0.0	0.1	0.5
												Average(abs)			0.125	0.13	0.221	0.135	0.144	0.207	1.088
												Median			0	0	0.125	0	0	0.063	0.594

Subject	Peer	1 (10%)	2 (20%)	3 (15%)	4 (25%)	5 (15%)	6 (15%)	Total Peer	Difference 1	Difference 2	Difference 3	Difference 4	Difference 5	Difference 6	Dif. peer	Difference self-peer	Levels for self	Levels for peer
1	3	100	0.25	0.30	0.25	0.19	0.38	**3.42**	0.125	0.000	-0.323	0.000	-0.313	0.250	**-0.45**	1.30	1/2	3
2	4	100	0.63	0.88	0.58	0.88	0.38	**6.90**	0.000	-0.313	0.047	-0.104	0.438	0.125	**0.03**	0.22	3	2/3
3	5																1	
4	6	0.50	0.75	0.75	0.50	0.50	0.75	**6.25**	-0.500	0.438	-0.063	-0.250	-0.125	0.188	**-0.25**	1.26	1	1
5	7	0.88	100	0.81	100	0.75	0.81	**8.94**	-0.125	0.063	0.031	0.000	-0.188	-0.125	**-0.42**	1.36	1/2	2/3
6	8	100	100	0.63	0.75	0.75	0.63	**7.88**	0.000	0.063	0.063	0.000	0.000	0.000	**0.22**	-0.97	1/2	2/3
7	9	100	0.75	0.75	0.50	0.50	0.75	**6.75**	0.000	0.000	0.109	-0.125	-0.250	0.000	**-0.52**	0.98	1/2	1
8	10	100	0.50	0.50	0.50	0.75	0.63	**6.06**	0.000	-0.125	-0.125	-0.250	-0.125	-0.063	**-1.34**	1.94	1	1
9	11	100	0.00	0.75	0.63	0.25	0.75	**5.19**	0.000	-0.938	0.219	-0.125	-0.625	-0.125	**-2.98**	4.96	1/2	2
10	12	100	100	0.75	0.25	0.69	0.63	**6.72**	0.000	0.063	0.125	-0.188	0.063	0.438	**0.59**	-1.75	1	1/2
11	13	100	100	0.63	0.88	0.75	0.75	**8.38**	0.000	0.000	-0.063	0.188	0.000	0.000	**0.38**	-1.42	1	1/2
12	1	100	100	0.85	0.75	0.75	100	**8.78**	0.000	0.250	0.104	-0.188	0.000	0.063	**0.28**	1.84	1	3
13	2	100	100	0.50	100	0.75	100	**8.88**	0.000	0.063	-0.281	0.000	-0.250	0.188	**-0.39**	2.39	1	1
						Average			-0.042	-0.036	-0.013	-0.087	-0.115	0.078	-0.4			
						Average(abs)			0.058	0.178	0.119	0.109	0.183	0.120	0.605			
						Median			0.000	0.031	0.039	-0.115	-0.125	0.031	-0.32			

5.5 Experiment 3

In the third experiment, third year students of a four year industrial engineering degree were required to model one part, as depicted in a detailed drawing (Figure 4.6). The model was explicitly required to allow for redesign so the height of the lateral wall would increase from 3 to 5 inches, and the width of the base would increase from 3.5 to 6 inches.

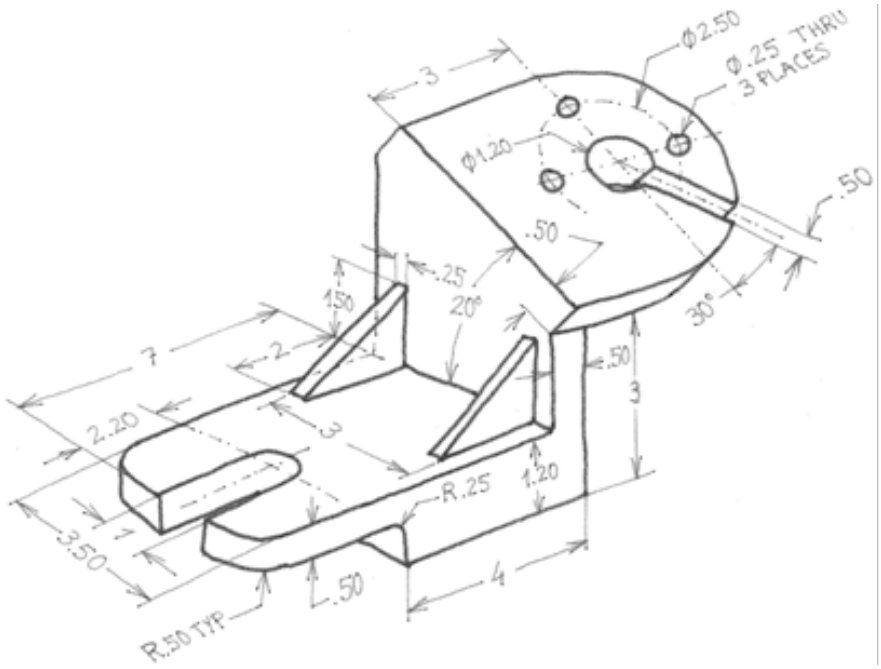


Figure 4.6. Detailed drawing of the part to model during Experiment 3.

Students had been incompletely introduced into rubrics from the beginning of the term (quality concepts had been explained by way of rubrics, but had never been applied). At the end of the exam,

they were given forty-eight hours to return the self-evaluation form. Then they were provided with paper forms and were informed that spreadsheet forms were also available at the Moodle-based learning platform [Moodle, 2015]. They all were also informed of a reward: their final marks would increase up to 10% if their self-evaluation equaled the instructor evaluation.

Two days after returning their self-evaluation, students were provided with a detailed explanation of the exam solution (including criteria to self-evaluate their work), and were informed that they could freely return a second self-evaluation under the following condition: only the better of the two self-evaluations would be used to calculate the 10% reward. This time, paper forms were not directly provided. Instead, they were notified that both paper forms and spreadsheet forms were available to download from the Moodle-based platform and could be used interchangeably.

The population of the study was forty-seven students, as forty-nine returned the exam, but two failed to return the rubrics. In the first evaluation, twenty-five students returned paper forms, and the other eight returned electronic files containing scanned images of the paper form. Eleven students returned spreadsheet forms and one student returned an unlocked spreadsheet form. Hence, thirty-three of forty-five students used paper forms while only twelve of forty-five used spreadsheet forms. However, in the second evaluation, only seventeen students returned rubrics (fifteen students returned spreadsheet forms). One unlocked spreadsheet form was also returned and only one student returned an electronic file containing the scanned image of the paper form.

Since paper forms were available for the first evaluation while spreadsheet forms had to be downloaded from the Moodle platform, it can be concluded that availability was the solitary reason for using paper forms. In fact, students found spreadsheet forms to be the best choice when both alternatives were equally available (both had to be downloaded from the Moodle platform).

5.6 Other results for Experiment 3

Detailed results for this experiment are shown in Table 4.3, which summarizes and compares differences between self and instructor evaluations for evaluations performed both before and after delivery of the exam solution. Differences are calculated for each of the six main dimensions of the rubric, with differences equal or higher than 50% highlighted in grey. All values are in the range [0, 1], except the values in bold, which are in the range [0, 10]. The right side of the table summarizes differences between evaluations performed before and after delivery of the solution.

Differences between instructor-evaluation and self-evaluation (in absolute value) for the forty-five students that returned the first rubric were calculated to be 1.23 in average and 0.96 in median. Differences between instructor-evaluation and self-evaluations for the seventeen students that returned the second rubric were calculated to be 0.68 in average and 0.50 in median. Two of the students that returned the second rubric failed to return the first rubric. The second rubric was the most accurate for ten of the fifteen students that returned both, while the first was the most accurate for the other five. These results suggest that students moderately improved their evaluations after having been exposed to the detailed explanation of the expected solution. Dimensions 1 and 2 required few modifications, while dimensions 3 to 6 appeared to have been poorly understood the first time. More precisely, up to 11 students modified their self-evaluation of Dimension 4, while only two modified Dimension 2.

Clearly, the results from this pilot experiment are not conclusive for various reasons, such as the fact that only students that believed they could improve over their first evaluation were motivated to return a second evaluation. Even so, introduction of rubrics during the exams instead of (or in addition to) introducing them along the teaching period, appears to be an alternative worth future consideration.

Table 4.3. Comparison between instructor, self, and peer evaluations for Experiment 3.

Subject	Difference 1	Difference 2	Difference 3	Difference 4	Difference 5	Difference 6	Difference Before	Difference Before (ABS)	Type of rubric returned	Difference 1	Difference 2	Difference 3	Difference 4	Difference 5	Difference 6	Difference After	Difference After (ABS)	Type of rubric returned	Best form	Before1-After1	Before2-After2	Before3-After3	Before4-After4	Before5-After5	Before6-After6	Before-After
1	-0.25	-0.63	-0.17	-0.25	0.50	0.75	**-1.38**	**1.38**	Paper																	
2	0.00	0.00	-0.17	0.38	0.13	0.25	**0.94**	**0.94**	Scanned paper	0.00	0.00	-0.17	0.63	0.38	0.00	**1.56**	**1.56**	Spreadsheet	Before	0.00	0.00	0.00	-0.25	-0.25	0.25	**-0.63**
3	0.00	0.00	-0.21	0.15	-0.13	-0.29	**-0.50**	**0.50**	Spreadsheet																	
4	0.00	0.00	0.25	0.13	0.25	-0.08	**0.92**	**0.92**	Paper																	
5	-0.25	-0.13	-0.08	0.75	0.50	0.58	**2.08**	**2.08**	Spreadsheet																	
6	0.00	0.00	0.17	0.13	0.00	0.58	**1.08**	**1.08**	Paper	-0.25	0.00	-0.08	0.13	-0.25	0.08	**-0.42**	**0.42**	Spreadsheet	After	0.25	0.00	0.25	0.00	0.25	0.50	**1.50**
7	-0.50	-0.75	0.08	0.50	0.63	0.50	**-0.19**	**0.19**	Paper																	
8	0.00	0.00	-0.08	0.38	0.13	0.00	**0.81**	**0.81**	Spreadsheet																	
9	0.00	-0.38	0.17	0.38	0.25	0.08	**0.33**	**0.33**	Paper																	
10	-0.25	0.00	-0.33	-0.13	0.00	0.33	**-0.67**	**0.67**	Scanned paper																	
11	-0.25	0.00	0.00	0.13	-0.25	0.33	**-0.04**	**0.04**	Paper																	
12	0.00	0.00	-0.25	0.38	0.50	0.00	**1.13**	**1.13**	Paper																	
13	0.00	0.00	0.00	0.13	0.25	0.33	**0.96**	**0.96**	Paper																	
14	0.00	0.00	0.08	-0.13	-0.25	0.25	**-0.25**	**0.25**	Spreadsheet																	
16	0.00	0.50	0.00	0.38	-0.25	0.83	**2.71**	**2.71**	Paper																	
17	0.00	0.00	-0.08	0.50	0.25	0.58	**1.83**	**1.83**	Paper	0.00	0.00	-0.08	0.25	0.25	0.33	**1.08**	**1.08**	Spreadsheet	After	0.00	0.00	0.00	0.25	0.00	0.25	**0.75**
18	0.00	0.00	-0.17	0.13	0.00	0.83	**0.83**	**0.83**	Spreadsheet	0.00	0.00	-0.17	-0.13	-0.50	0.58	**-0.67**	**0.67**	Spreadsheet	After	0.00	0.00	0.00	0.25	0.50	0.25	**1.50**
19	0.00	-0.25	0.08	0.38	0.50	0.33	**1.21**	**1.21**	Spreadsheet unlocked	-0.50	-0.25	0.08	0.13	0.50	0.33	**0.21**	**0.21**	Spreadsheet	After	0.50	0.00	0.00	0.25	0.00	0.00	**1.00**
20	0.00	0.13	0.25	0.38	0.25	0.67	**2.54**	**2.54**	Paper																	
21										-0.50	-0.25	0.33	0.63	0.50	0.58	**1.83**	**1.83**	Spreadsheet unlocked								
22	0.00	0.00	-0.25	0.00	0.00	0.17	**-0.21**	**0.21**	Spreadsheet																	
23	0.00	0.00	-0.25	0.38	0.13	-0.17	**0.40**	**0.40**	Paper																	
24	0.00	0.00	0.25	0.25	0.00	-0.08	**0.79**	**0.79**	Paper	0.00	0.00	-0.04	0.33	-0.06	-0.08	**0.43**	**0.43**	Spreadsheet	After	0.00	0.00	0.29	-0.08	0.06	0.00	**0.36**
25	0.00	0.00	-0.17	0.50	-0.25	0.00	**0.38**	**0.38**	Paper	0.00	0.00	0.08	0.44	-0.13	0.13	**0.94**	**0.94**	Spreadsheet	Before	0.00	0.00	-0.25	0.06	-0.13	-0.13	**-0.56**
26	0.00	-0.13	-0.17	0.50	0.00	0.25	**0.63**	**0.63**	Spreadsheet																	
27	0.00	0.00	0.25	0.38	0.25	0.33	**1.83**	**1.83**	Paper																	
28	0.00	0.00	0.17	0.75	0.50	0.50	**3.00**	**3.00**	Scanned paper																	
29	0.00	0.13	0.08	0.38	0.25	0.42	**2.04**	**2.04**	Paper																	
30	0.00	0.13	0.00	0.38	-0.25	0.00	**0.75**	**0.75**	Paper	0.00	0.13	0.00	0.38	-0.25	-0.25	**0.50**	**0.50**	Spreadsheet	After	0.00	0.00	0.00	0.00	0.00	0.25	**0.25**
31	0.00	0.00	0.00	0.13	-0.25	0.17	**0.04**	**0.04**	Scanned paper	0.00	0.00	0.25	0.13	-0.25	0.17	**0.42**	**0.42**	Spreadsheet	Before	0.00	0.00	-0.25	0.00	0.00	0.00	**-0.38**
32	0.00	0.00	0.50	0.25	0.25	0.50	**2.13**	**2.13**	Spreadsheet																	
33	-0.25	-0.38	0.08	0.38	0.75	0.42	**1.04**	**1.04**	Scanned paper																	
34	0.00	0.00	0.17	0.75	0.00	0.08	**1.83**	**1.83**	Scanned paper	0.00	0.00	-0.01	0.21	-0.06	0.04	**0.34**	**0.34**	Spreadsheet	After	0.00	0.00	0.18	0.54	0.06	0.04	**1.49**
35	0.00	0.13	0.25	0.75	0.88	0.42	**3.98**	**3.98**	Spreadsheet																	
37	0.00	-0.25	0.08	0.50	-0.25	0.17	**0.17**	**0.17**	Paper	-0.25	0.00	-0.17	0.25	-0.25	0.17	**-0.21**	**0.21**	Spreadsheet	Before	0.25	-0.25	0.25	0.25	0.00	0.00	**0.38**
38	0.00	0.00	0.25	0.38	0.25	-0.08	**1.42**	**1.42**	Paper																	
39	0.00	0.00	0.58	0.13	0.38	0.50	**2.19**	**2.19**	Paper																	
40	0.00	0.00	-0.17	0.13	0.00	-0.08	**-0.08**	**0.08**	Scanned paper																	
41	0.00	0.13	0.33	0.25	-0.06	0.17	**1.45**	**1.45**	Spreadsheet	0.00	0.13	0.08	0.13	0.00	0.08	**0.83**	**0.83**	Spreadsheet	After	0.00	0.00	0.25	0.13	-0.06	0.08	**0.61**
42	-0.25	-0.13	-0.08	0.00	0.00	0.17	**-0.58**	**0.58**	Paper																	
43	0.00	0.00	0.25	0.38	0.88	0.83	**3.27**	**3.27**	Paper																	
44	-0.25	0.00	0.00	0.13	0.50	0.25	**1.00**	**1.00**	Paper																	
45	0.00	0.00	0.17	-0.25	0.00	0.25	**0.00**	**0.00**	Paper	0.00	0.00	-0.08	-0.25	-0.25	0.25	**-0.75**	**0.75**	Spreadsheet	Before	0.00	0.00	0.25	0.00	0.25	0.00	**0.75**
46	0.00	0.00	0.33	0.13	0.00	0.00	**0.75**	**0.75**	Scanned paper	0.00	0.00	0.08	-0.13	0.00	0.00	**-0.13**	**0.13**	Scanned paper	After	0.00	0.00	0.25	0.25	0.00	0.00	**0.88**
47	-0.13	0.00	0.19	0.69	0.38	0.58	**2.69**	**2.69**	Spreadsheet	-0.19	-0.25	-0.08	0.48	0.25	0.54	**0.81**	**0.81**	Spreadsheet	After	0.06	0.25	0.28	0.21	0.13	0.04	**1.88**
48										0.00	0.00	-0.08	-0.13	-0.25	0.25	**-0.50**	**0.50**	Spreadsheet								
49	0.00	0.13	0.25	0.38	0.50	0.25	**2.50**	**2.50**	Paper																	
																			Changes	4	2	10	11	9	9	
			Average				1.06	1.23					Average			0.37	0.68	Before	5	Average	0.07	0.00	0.10	0.12	0.05	0.10	0.65
			Median				0.94	0.96					Median			0.42	0.50	After	10	Median	0.00	0.00	0.18	0.13	0.00	0.04	0.75

6. Discussion

First of all, it can be concluded that it is feasible to use adaptable rubrics, which provide immediate feedback of the evaluation:

- No problems were reported while using adaptable rubrics in the experiments, apart from a single student that manipulated the internal parameters of the spreadsheet form. However, all

instructors reported a large workload in preparing the spreadsheet forms and extracting the information from the submitted forms. Hence, adaptable rubrics are operative, but this tool requires improvement, since spreadsheet forms are unsecure and impractical.

- The notion that students could be replying randomly or fraudulently can be dismissed, since student evaluation results were quite similar to the instructors (not only their own task, but also those performed by other students). Thus, it can be concluded that the students appear to understand the majority of the rubric and feel comfortable with the electronic form.
- Since there is great dispersion in the use of levels, it can also be concluded that allowing the use of different levels on demand is clearly helpful.
- In principle, score feedback appears to also be useful, since all undergraduate and master's students replied to the question of whether "Feedback on qualifications is helpful" with answers "Nearly always" and "Yes/Always." However, as is discussed shortly, the credibility of this determination is slightly dubious.
- It is assumed that both formative assessment and self-learning would be easier with adaptable rubrics. Hence, it is worth performing a full-scale experiment to validate or reject the hypothesis if adaptable rubrics assist students in better self-evaluation of their own task.

As shown in Figure 4.1, a short satisfaction test was included in the spreadsheet form. However, the main conclusion from the experiment with undergraduate students is that students appear to reply more optimistically than accurately (perhaps in erroneous belief that this may increase their possibility to get a better mark). The question, "I understand the criteria" was replied with median responses as "Nearly always". The same occurred for the question, "Feedback on qualifications is helpful". Answers were still more optimistic (in between "Nearly always" and "Yes/Always") for questions Two and Three: "Sub-criteria are helpful" and "This e-form is better than paper". However, this response is a clear contradiction, as eleven of thirty students who claimed that sub-criteria are helpful, only used Level 1. Similarly, master's students answered between "Nearly always" and "Yes/Always" to the question, "This e-form is better than paper" although they previously had never used the paper form. Hence, instead of direct questions that are not necessarily truthfully answered, possibly it would be better to allow students the freedom to choose between paper and spreadsheet forms so as to determine their actual preferences.

During the experiments, some students declared that some criteria were hard to understand (in clear contradiction with replies to Question 1 of the satisfaction test). Furthermore, the instructors agreed that using the rubric criteria was difficult, so a decision was made that the assertions should not define the goal, since it is usually an abstract concept for the student to comprehend. Instead, they should describe what is checked for the evaluation of each criterion. For instance, instead of asking whether the "constraints in the profiles emphasize the purpose of the model" (Criterion 6.1.a), it is suggested to first distinguish between geometrical and dimensional constraints. The criterion should then be restated as whether "dimensions in the model's profiles replicate dimensions of the input, or if transferred, whether the new ones follow a clearly understandable purpose". In other words, instead of introducing quality concepts top-down, the criteria should be introduced bottom-up.

Finally, an improved electronic form was deemed necessary, and will be discussed in the next section. It has been stated that spreadsheet rubrics are not portable, as they require resources such as macros that are blocked by most anti-virus applications. Besides, resources such as ActiveX controls must be loaded, which results in delays and "screen blinking" while the forms are opened. Another important problem is the relative difficulty encountered in processing the information of the submitted rubrics. Finally, spreadsheet rubrics can be unlocked easily, allowing for inconsistent scores (i.e. differently scoring the main criterion for each dimension and its sub-criteria). Preventing those inconsistencies (quite common in paper forms), was one goal for developing electronic rubrics, and spreadsheet forms had proven not valid for that purpose. Preparing the prototype rubric required expertise in spreadsheets

and extracting information from the spreadsheet rubrics was time consuming (roughly one workday for every experiment) while also being prone to error.

While this situation is acceptable for a pilot experiment (were the goal was validating new ideas and preparing a full scale experiment to validate or reject them), it would not be acceptable during normal teaching and learning processes. Hence, the derived final e-rubric needed to be implemented to adapt easily to different evaluation needs while also integrating into complex teaching scenarios supported by current teaching platforms.

7. Preliminary Conclusions

An adaptable rubric prototype using spreadsheets was developed as a proof of concept to test adaptable rubrics, as achieved with a tree structure around different dimensions (criteria) which are recursively developed in increasing levels of detail. The users are allowed to freely scroll or unscroll details of the criteria of interest and the rubric simultaneously provides immediate feedback of the evaluation scores.

The experiments conducted illustrate that spreadsheet rubrics are useful to verify the utility of the adaptable rubric specifications and their advantages over paper form rubrics. Conversely, they have also revealed certain disadvantages: preparing spreadsheet rubrics requires expertise in spreadsheet programming and extracting information from the spreadsheet rubrics is time consuming. They provide feedback of scores but do not allow for automatic feedback on marking criteria, which is suggested by some researchers [Nield et al., 2014]. The detected weaknesses focus on the next steps toward a more efficient adaptable rubric. The next development was replacing spreadsheets with a web-based framework to manage rubrics and users. The next task was designing, implementing, and testing “adaptable e-rubrics”. Ideally, this tool for rubrics could be used as a stand-alone device or integrated in the teaching/learning flow as a plug-in of a learning management system (LMS).

The experiments described above were performed by a research team of which I was a member. My specific responsibilities included data assimilation, reviewing relevant literature, and manuscript co-authoring.

8. e-Rubrics Overview

The elaborated criteria used to assess MCAD model quality require a more specialized computer supported rubric tool, as current paper and spreadsheet-based formative rubrics, as well as Computer Assisted Assessment (CAA) systems are inadequate to properly convey CAD model quality [Company et al., 2015].

The concept of adaptable rubrics emerged as a powerful mechanism to support different learning styles and rhythms. Adaptable rubrics provide multiple levels of detail that can be expanded on demand. Therefore, the level of detail can be adjusted and adapted to a specific teaching scenario and/or the students’ level of understanding of quality concepts. This adaptable capability allows the use of rubrics as a formative tool to convey MCAD quality criteria that can be introduced at the beginning of the training period and integrated throughout the entire instructional process.

A new computer-assisted rubric platform is presented, designed specifically to support adaptable rubrics. This platform provides feedback (showing detailed scores and levels of performance, if requested), adapts to the learning rhythm and style of each student (different levels of detail can be deployed on demand, according to student needs), and also collects metadata. The platform is completely generic, as it can be used to manage different rubric types. Implementation strategy, validation, and lessons learned while developing and testing this platform to convey concepts about quality to new CAD trainees are the core of this work.

The remainder of the chapter is structured as follows:

- Section 9 describes the state of the art in platforms for scoring rubrics and confirms the lack of support for adaptable rubrics.
- Section 10 describes the architecture of the proposed system, detailing design specifications and the most relevant implementation details.
- Section 11 describes three experiments aimed at validating the system. The first experiment validates the hypothesis that the new platform does not negatively affect the reliability of the rubrics submitted by the students. The second experiment shows that the new platform provides richer and more meaningful information than other systems. The third experiment demonstrates an approach for introducing adaptive rubrics by taking advantage of the synergistic effect of combining suitable lecture notes during instruction and adaptable rubrics for feedback and evaluation.
- Section 12 concludes by highlighting the lessons learned which will be used to guide future developments and improvements of the platform.

9. Supplementary Background

Many authors claim that rubrics can both teach and evaluate [Andrade and Du, 2005], [Jonsson and Svingby, 2007], [Malini and Andrade, 2010]. In fact, the formative use of scoring rubrics has proven useful under many circumstances [Panadero and Jonsson, 2013]. Unfortunately, the use of rubrics in the context of MCAD teaching/learning is a non-trivial challenge. MCAD systems address part of a complex problem [Carberry and McKenna, 2014] and are complex software tools that require not only a thorough understanding of the various functionalities provided by the system, but also the application of efficient strategies to create high quality CAD models. Based on related research, this issue can be considered a particular type of the "content-based constructed responses" problem [Liu et al., 2014].

Computer-assisted grading rubrics are essential to improve the efficiency and effectiveness of grading [Anglin et al., 2008] and [Auvinen et al., 2009]. Although some commercially available CAA tools provide automatically contextualized feedback [Santos et al., 2009] and [Cebrián-Robles et al., 2014], and some Learning Management Systems (LMS) [Atkinson and Lim, 2013] provide rubric functionality, currently available implementation only support static rubrics that does not provide any flexibility to adapt to different learning scenarios.

Building on the general concepts of using rubrics to capture judgment [Mertler, 2001] and [Karkehabadi, 2013] and the evaluative use of rubrics in higher education [Malini and Andrade, 2010], this ambit focuses on the role of rubrics as instruments for acquiring or reinforcing complex skills [Manson and Olsen, 2010] and [Smit and Birri, 2014], assessing engineering questions [McCormick et al., 2015], and supporting the process of formative assessment [Pophan, 1997] by advising students about their progress and assisting them in their development [Panadero and Jonsson, 2013].

Contrary to holistic rubrics (which only score the overall process or product as a whole, without judging the individual components separately [Nitko, 2001]) analytic rubrics are typically used for formative feedback [Mertler, 2001], as they allow the individual scoring of parts (or dimensions) of the product or performance [Moskal, 2000]. Of particular interest is general rubrics, as they provide better formative assessment than task-specific ones (where their initial low reliability gradually improves as students get accustomed to the rubrics [Brookhart, 2013]).

Nonetheless, rubrics should be accompanied by suitable lecture notes. It has been argued that providing study guides with learning objectives that deal with all aspects of the problem solving process and administering individual tests consistent with the study guides constitute a good

instructional method, at least in addressing Outcome 3-c of the program educational objectives of the ABET engineering criteria for courses in which design problems are assigned [Felder and Brent, 2003].

A common problem when introducing performance assessment is measuring complex competences in a credible way (i.e., whether or not observations of complex behavior can be performed in a reliable and trustworthy manner) [Jonsson and Svingby, 2007]. Therefore, the first problem is determining how well rubrics convey MCAD quality criteria.

When comparing instructor and student judgments, assessment differences can be attributed to the fact that the concepts under evaluation may still be so foreign that students are unable to recognize them [Orsmond et al., 1996]. This lack of consensus between the instructor assessment and the student self-assessment (i.e., the lack of inter-rater reliability) is useful to detect problems in the understanding of quality criteria. This type of data can be processed by applying proper statistical analyses to the assessed rubrics [Zaiontz, 2015].

Additionally, metadata that describes how users interact with the platform while completing the rubrics is also of interest. It has been argued that rubrics should be complemented with "anchors" (i.e. written descriptions), examples that illustrate the various levels of attainment, or work samples [Jonsson and Svingby, 2007]. No results have been reported regarding the use of metadata to find improved methods to increase the effectiveness of anchors.

Although advanced trainees may not require detailed explanations of each evaluation criterion, more detail may be desirable for novice users. Adaptable rubrics should allow students to display more detail and score low level criteria on demand. If a particular criterion in the rubric is too abstract or difficult to understand, the additional level of detail can provide a clearer description of the expected performance levels [Company et al., 2015]. A major challenge with adaptable rubrics involves guaranteeing consistent scoring for each different combination of detail-level responses. In this arena, spreadsheets have proven impractical (as shown earlier in this chapter), as its implementation requires significant programming proficiency and extracting relevant information from them is time consuming. Furthermore, they can hardly convey performance criteria descriptors, which are critical components of rubric design [Tierney and Simon, 2004].

10. Electronic Platform for Adaptable Rubrics

This section describes the design and implementation of a dedicated rubrics platform, with current support for adaptable and future support for adaptive rubrics, and the ability to output metadata to analyze the evaluation process to guide further improvements.

10.1 Design Specifications

The main goal of the design was to centralize both rubrics and assessment results in order to improve data exchange while reducing the time required to program and manipulate a rubric. An additional goal was to simplify the rubric creation process by providing specialized and intuitive tools to enter new rubrics into the system and edit existing ones. Finally, metadata management to support learning analytics was also included as an expected functionality. Metadata that describes the context of the evaluation process is required to better understand current problems and guide future improvements. Metadata is defined as the dataset collected by tracking how users interact with the platform while filling out the rubrics. Spreadsheet forms (which cannot collect such data) and proprietary "black-box" rubric platforms (which do not grant direct access to the data) are unsuitable for this goal.

Two main specifications were defined for the rubrics platform: (1) the system must be adaptable allowing for rubrics with varying levels of detail, and (2) the rubrics should be easy to reconfigure, allowing instructors the capability to adjust them to changing scenarios.

The first specification implies that every participant that uses the rubric to score a particular task should be able to select the desired level of detail. Three levels of detail are required due to the structure of the quality of MCAD models that are employed, which distinguishes between main criteria (Level 1) and two nested levels of sub-criteria (Levels 2 and 3). Level 1 is the highest level of abstraction, as it directly evaluates the six main dimensions associated with quality of CAD modeling defined by [Company et al., 2015]: validity, completeness, consistency, conciseness, clarity, and design intent. Levels 2 and 3 progressively give more details for each particular dimension. The second specification is intended to allow instructors to customize rubrics based on the evolution of student learning. Long term goals include linking the rubrics with appropriate anchors (as defined by [Jonsson and Svingby, 2007]).

Three auxiliary design specifications were also considered: (3) the system must provide immediate feedback of the evaluation scores; (4) instructors should be able to easily extract and process the information of the completed forms, and (5) the forms should prevent incomplete or inconsistent scoring.

From a user standpoint, the platform manages three types of accounts: students, instructors, and administrators. After a successful login to the system, students are allowed to perform evaluations using rubrics (both self and peer evaluation). Instructors are allowed to manage rubrics, student users and groups, assign rubrics to groups, schedule rubric assignments, and manage results. New instructors can be created by a global administrator.

10.2 Implementation

As part of this work, the research team developed Annota eRubrics, a web based framework to manage rubrics and users. Three user profiles were defined: student, instructor, and global administrator, as discussed in the previous section. Student users can fill out rubrics and visualize completed rubrics during the evaluation period defined by the instructor. Instructors can manage rubrics (create new, import from CSV or XML files, export to files, and assign rubrics to students during a certain time period), manage students (register new, import from CSV files, and manage groups of students), and access assessment results (view and export to files). Finally, global administrators (who can also be instructors) are able to register new instructors.

The general structure of the proposed framework is shown in Figure 4.7. Data is stored in a MySQL database and accessed via an Apache web server by a number of PHP scripts triggered by the framework front-end running over Unity3d Webplayer [Creighton, 2010]. Only a name and an e-mail address is required to create a new user. During registration, a custom link is automatically emailed to the users, so they can set up their password and enter basic demographics information to complete the process. A screen capture of an assembly rubric is shown in Figure 4.8.

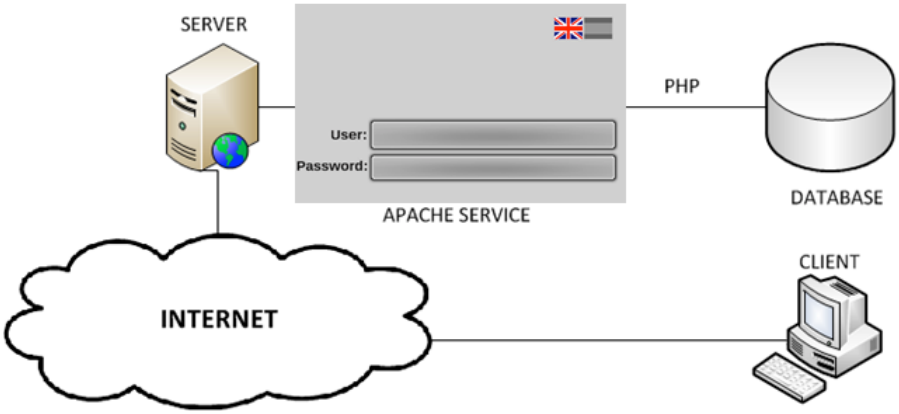


Figure 4.7. Annota framework.

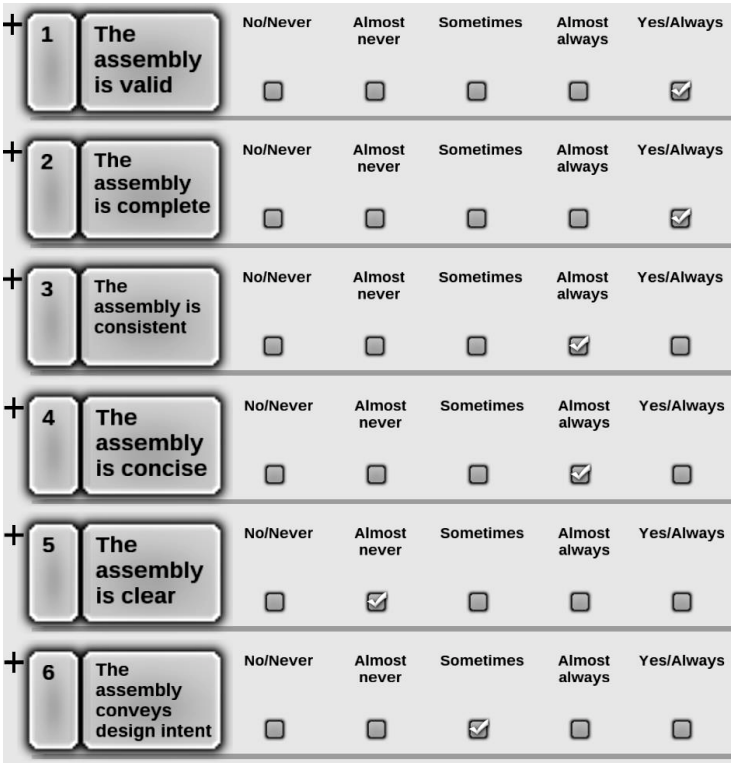


Figure 4.8. Assembly rubric as available in the Annota platform.

An example of the mechanism to unfold low level sub-criteria is illustrated in Figure 4.9. By clicking the "+" (unfold) or "–" (fold) signs located on the upper left corner of the criteria (Criteria 1 and 1.2 have been unfolded in Figure 4.9), users can dynamically adapt the rubric to their own rhythms and learning styles. Instructors can pre-configure criteria as "folded" or "unfolded" to determine which criteria will be folded/unfolded by default when the student first loads the rubric.

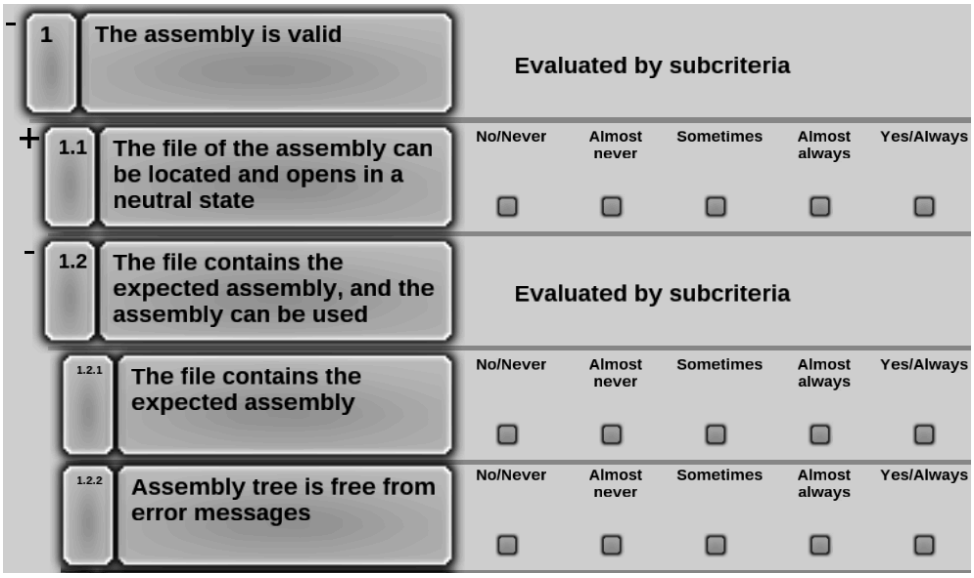


Figure 4.9. Annota rubrics are adaptable by allowing users to fold and unfold levels of detail.

Each criterion has an associated scoring weight configured by the instructor. Assessment results are recalculated every time a criterion is filled out by the student. By default, weights and assessment results are not visible to the student, but can be activated at any time. This activation/deactivation is part of the metadata collected by the system.

Figure 4.10. Criterion with an associated scoring weight.

As a result of the first experiment (described in the next section), contextual information bubbles were added to various options of the rubric. These bubbles display textual explanations for each level of deployment, also called performance criteria descriptors, and are automatically displayed when the user hovers over a check box (Figure 4.11). The number of bubble activations and the bubble activation times are also part of the collected metadata.

Figure 4.11. Example of bubble with textual explanation for a particular level of deployment.

From an instructor standpoint, a "student view" is available, which allows the instructor to examine the rubric as a member of the student group. This mode is used for testing purposes only, so results are not saved. In addition, instructors can grade students by importing a csv file. This functionality allows the use of spreadsheets to grade a framework user group with a single mouse click.

11. Experiments

To validate the platform, three experiments were conducted with various groups of CAD users. The first experiment demonstrates that the new rubrics platform does not negatively affect the reliability of the evaluations. The second experiment illustrates how the metadata generated by the new platform can be used as a resource to identify current weaknesses, measure the effectiveness, and lead to further improvements. The third experiment validates this approach to CAD training by combining suitable tutorials and materials provided during instruction and a set of adaptable rubrics for feedback and evaluation.

11.1 Experiment 1

To determine whether reliability is affected by the form used in the assessment, three scenarios where students used paper, spreadsheets, and Annota forms were compared. Because using different populations would contaminate the experiment with uncontrolled differences, and repeatedly asking the same population to re-evaluate similar work (although with different forms) would result in invalid comparison (since students gradually improve their reliability as a result of their exposure to previous evaluations), it was decided to use the same group of subjects, and ask them to evaluate different works. First, they evaluated a single part model; next, an assembly; and finally a 2D working drawing.

The sample included a group of junior industrial engineering students from a Spanish university. All students had basic knowledge of rubrics (CAD quality concepts were explained using rubrics, but they were never put into practice). The total sample size was forty-nine students, but only thirty-five participants submitted the evaluations. Completion of the evaluations was presented as an optional (but recommended) part of the modeling task. Participants were encouraged and rewarded for submitting their evaluations.

Initially, students were asked to create a 3D solid model of a part depicted in a detailed drawing (see Figure 4.12). Reusability was explicitly stated as a major requirement for the model. Students were asked to follow an efficient strategy and create a flexible and reusable model so the part could be easily redesigned. The requested design changes involved increasing the height of the lateral wall from 3 to 5 inches, and the width of the base from 3.5 to 6 inches.

After modeling the part, students were asked to self-evaluate their work using rubrics in either paper or spreadsheet forms. Thirteen subjects were allowed to evaluate their models a second time because they admitted that they disliked their initial evaluations, which were done in a short period time right after the modeling task. At the end, eighteen students completed paper forms and the remaining seventeen submitted spreadsheet forms.

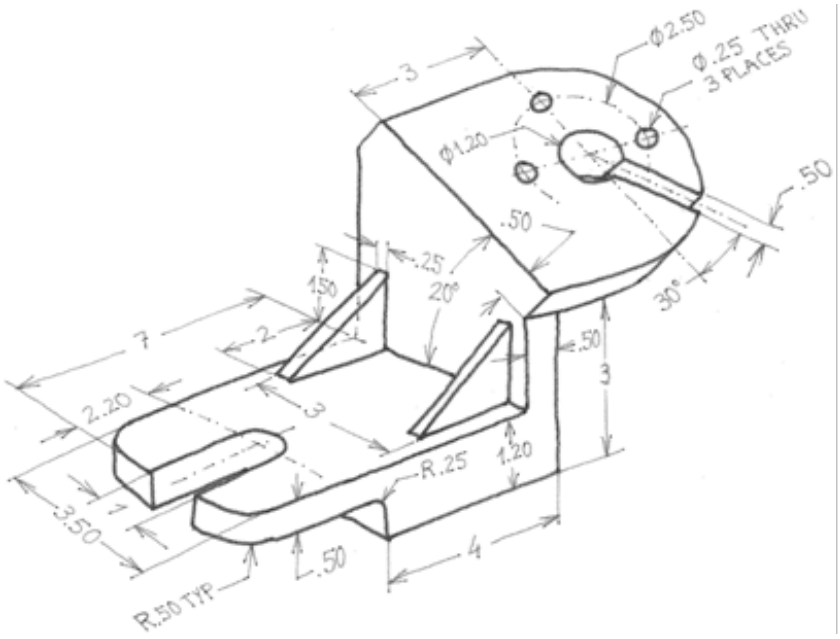


Figure 4.12. Detailed drawing of the part used in Experiment 1.

In order to determine whether students understood the CAD quality criteria introduced by the instructor, the reliability of the students' evaluations were measured, using the instructor's evaluation as the ideal. Thus, inter-rater differences were determined (between the self-evaluations and the instructor evaluations) for the six dimensions (see Figure 4.2) of the quality criteria [Company et al., 2015].

Annota rubrics use a 5-point Likert scale and internally map the evaluations into a numerical scale ranging from 0 to 1 (i.e. 0, 0.25, 0.5, 0.75, and 1, respectively from No/Never to Yes/Always). It is assumed that either the instructor or the students may evaluate using only the higher level criteria, thus a difference of less than 0.25 between the student and the ideal represented by the instructor evaluation is labeled as "agreed" and used to calculate the percentages of agreement shown in Table 4.4. The reliability (measured by the percentages of agreement) slightly increases from the first to the second rubric. This finding is expected, because of the accumulated experience of the students that completed the self-evaluations twice. The same effect is noticeable if the rubrics are disaggregated in paper forms from those that used spreadsheet forms, since the latter were filled out mainly during the second evaluation round. By analyzing the percentages of agreement for the entire group ("all subjects" row), it can be concluded that only the two most basic dimensions (1 and 2) appear to have been assimilated satisfactorily by the students (percentages of agreement ≥ 0.75).

Important differences in the assimilation process of the six dimensions are evident, which suggests that the anchors described previously might be necessary improvements for successfully introducing quality concepts in CAD instruction. Differences are due to the fact that uninformed students still tend to "frame" the problem of obtaining a 3D model, thus ignoring that the model is part of a "problem

solving" strategy. Hence, there is a need to improve the lecture notes to emphasize that CAD models must contribute to "delay making design decisions in order to explore, comprehend and frame the problem better" [Crismond and Adams, 2012].

Table 4.4. Percentages of agreement between self-evaluations and instructor evaluation for modeling task.

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
First rubric	73	91	59	41	32	32	55
Second rubric	85	92	92	38	23	54	64
Paper forms	67	89	50	39	22	33	50
Spreadsheet forms	88	94	94	41	35	47	67
All subjects	77	91	71	40	29	40	58

The next part of the experiment involved two tasks. Participants were asked to (a) assemble a virtual device (a type of pneumatic cylinder) provided with all the components, and (b) create the assembly drawing (including part numbers and bill of materials). The assemblies are shown in Figure 4.13. The components comprising the spring return cylinder (Figure 4.13-left) were provided to a first sub-group of fourteen students (who took the test first), while the double-acting cylinder (Figure 4.13-right) was the task for the second sub-group of twenty-three students (who were tested one day later).

For this activity, students were given 48 hours after the exam to submit the self-evaluation. They were granted access to the Annota rubrics platform and offered a completion reward of up to 10% extra credit, based on how close their self-evaluation was to the instructor's evaluation.

The aggregated differences for the entire group and the differences for each particular sub-group are shown in Tables 4.5 and 4.6. As mentioned earlier, all participants had basic exposure to rubrics (as they all used paper or spreadsheet forms to evaluate their performance while modeling), but they had no specific background on rubrics for measuring performance in assembly modeling or extracting drawings from assemblies.

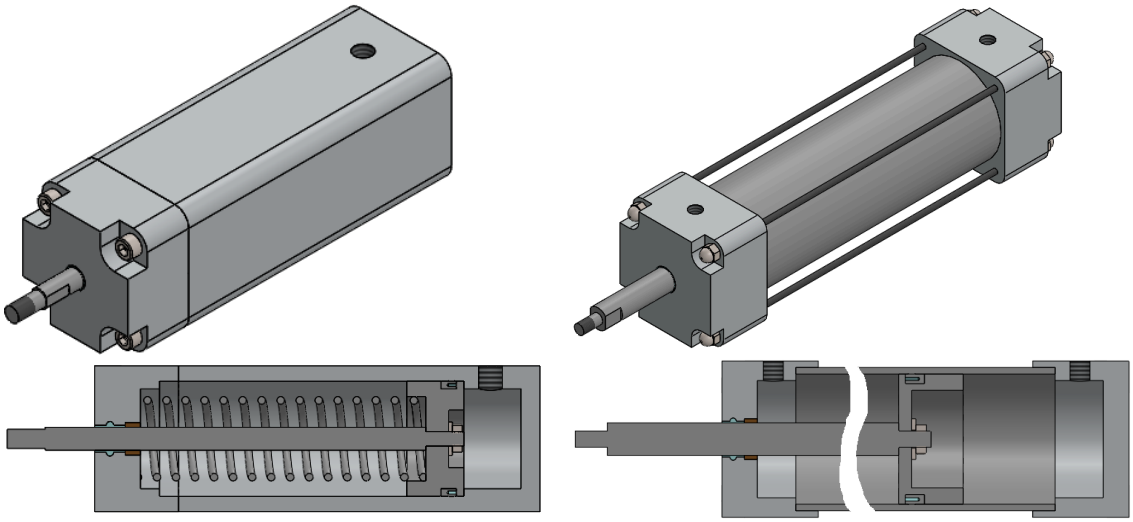


Figure 4.13. Pneumatic cylinders used in Experiment 1.

Because the second sub-group was examined 24 hours after the first sub-group, some participants had prior knowledge about the particular rubrics being used. Nevertheless, only a slightly significant difference was observed, (in favor of sub-group 1). A probable cause (not further investigated) could

be the lower level of instructor involvement in the second sub-group in terms of introducing quality concepts through rubrics.

Table 4.5. Percentages of agreement between self-evaluations and instructor evaluation for assembly task.

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
First sub-group	66.7	100	83.3	75	25	16.7	**61**
Second sub-group	60.9	87	47.8	52.2	39.1	4.35	**49**
All subjects	62.9	91.4	60	60	34.3	8.57	**53**

Table 4.6. Percentages of agreement between self-evaluations and instructor evaluation for drawing task.

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
First sub-group	58.3	66.7	66.7	50	66.7	66.7	**63**
Second sub-group	82.6	47.8	30.4	60.9	65.2	56.5	**57**
All subjects	74.3	54.3	42.9	57.1	65.7	60	**59**

By comparing Tables 4.5, 4.6, and 4.7, it can be observed that there are no significant differences in the inter-rater scores. Average percentages of agreement remain in the 53% to 59% range. In addition, Crombach's alpha was also calculated, although only alphas for the full group ("all subjects") were compared, since distinguishing sub-groups results in small sample sizes resulting in negative alphas. The result indicates that alpha is consistently low, for both paper-spreadsheet (α= 0.29) and Annota rubrics (α= 0.26 for the assembly and α= 0.29 for the drawing), as expected for a multidimensional structure. Therefore, it can be concluded that the new electronic rubrics platform (Annota) does not negatively affect the reliability of the evaluations.

11.2 Experiment 2

The second experiment illustrates that the new platform provides richer and more complete information, as it outputs user interaction metadata that can be used to analyze the evaluation process and guide to further improvements.

The same sample used in Experiment 1 was also used in this experiment. In this instance however, all forty-seven students were included, as all participants submitted valid rubrics and no significant differences were found between the thirty-five students that completed the first experiment and the remaining twelve students.

First, students were required to create a solid model of a specific part (a fixed arm of a pistol clamp) depicted in a detailed drawing (see Figure 4.14). Then, they were asked to create their own detailed drawing of the fixed arm they previously modeled. Finally, participants were required to assemble a virtual model of the pistol clamp (see Figure 4.15). All parts were provided with the exception of the standard parts and the fixed arm.

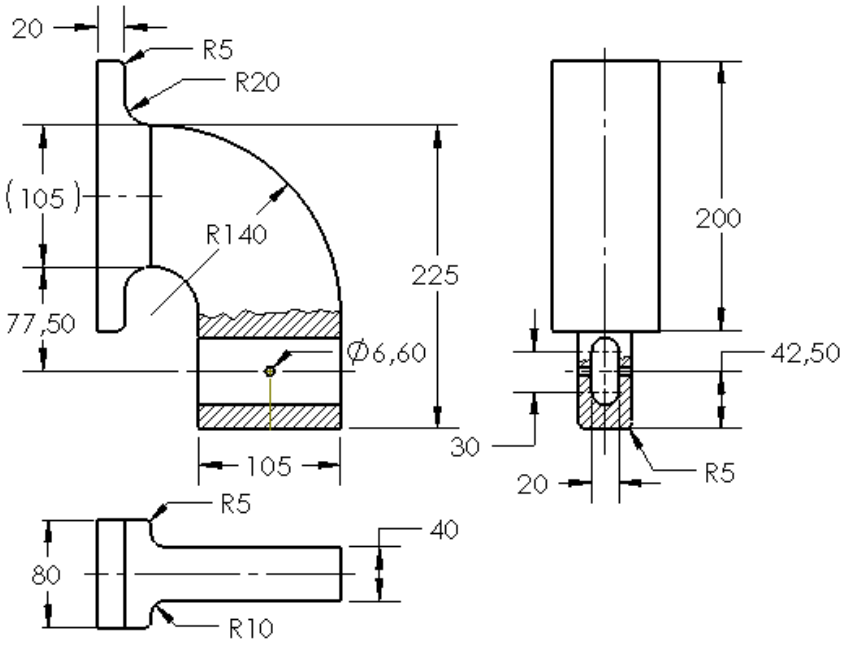


Figure 4.14. Detailed drawing of the part used in Experiment 2.

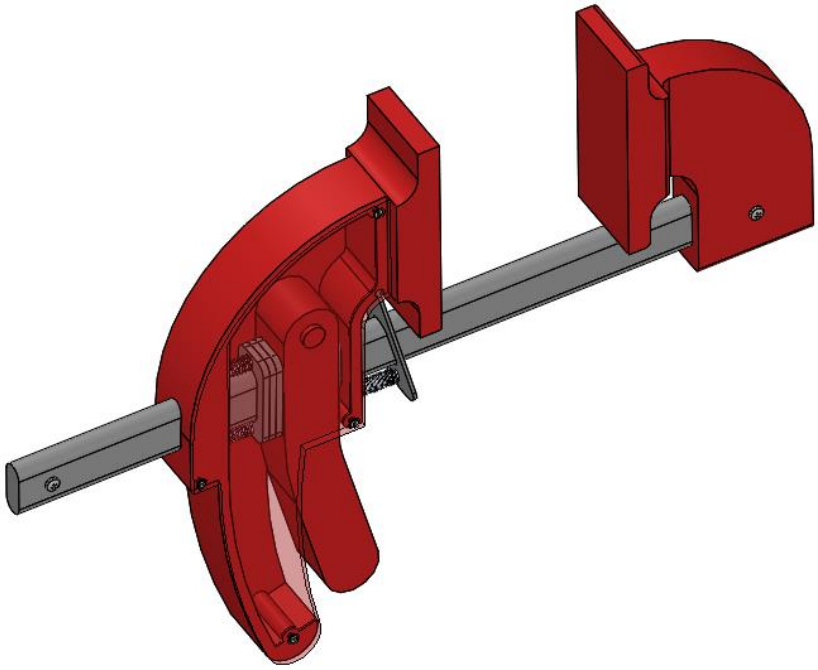


Figure 4.15. Axonometric view of the pistol clamp used in Experiment 2.

Percentages of agreement (see Table 4.7) confirm that the reliability of the evaluations gradually increases (because of the accumulated experience of the students), and only the two most basic dimensions (1 and 2) appear to have been assimilated well enough by the students. Nevertheless, certain aspects of the tasks seem to have been missed by a number of students.

Table 4.7. Percentages of agreement between self-evaluations and instructor evaluation for Experiment 2.

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
Modeling task	91	68	51	32	32	9	47
Drawing task	77	68	43	51	21	17	46
Assembling task	77	83	40	40	57	30	55

As a result of the first experiment, Annota was improved by incorporating rectangular bubbles that contextually display textual explanations for each level of deployment (as shown previously). In addition, the new platform collects information about how users interact with it while filling out the rubrics. In particular, the platform tracks how long each detailed explanation was visible, so this data can be correlated with the percentage of agreement of each student.

For the modeling task, it can be concluded that students that looked at the bubbles longer than average have a percentage of agreement that is generally greater than the rest (Table 4.8). Therefore, using bubbles to provide explanations seems to be a valid anchor to illustrate the various levels of attainment.

Table 4.8. Percentages of agreement between the students using bubbles and the rest.

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
More bubbles than average	100	71	65	29	41	18	54
Modeling task (all students)	91	68	51	32	32	9	47
Less bubbles than average	87	67	43	33	27	3	43
More bubbles than average	75	75	50	63	25	31	53
Drawing task (all students)	77	68	43	51	21	17	46
Less bubbles than average	77	65	39	45	19	10	42
More bubbles than average	70	70	30	40	70	20	50
Assembling task (all students)	77	83	40	40	57	30	55
Less bubbles than average	78	86	43	41	54	32	56

Additionally, the capability of the platform to collect metrics based on user interaction clearly assists in finding better strategies to convey CAD quality concepts through rubrics. As an example, an examination of the metadata about the use of bubbles reveals that the bubbles linked to Dimension 1 were helpful, as the average time was greater for Dimension 1 (Table 4.9), while the percentage of agreement for this dimension was very high (up to 100% in Table 4.8). However, Table 4.9 also reflects a much activity using bubbles for Dimension 6 (the high number of students that looked at the explanations linked to the sixth dimension, the longest time was used for Dimension 6, etc.), but it seems that having this extra information was not sufficient, as the percentage of agreement for this dimension continued to be extremely low. It is speculated that the explanations for each level of deployment need improvement (at least for Dimensions 6, and possibly also for Dimensions 4 and 5).

The duration during which bubbles were displayed in the assembly and drawing rubrics (Table 4.8) tell an entirely different story. Students that used bubbles agreed less (on average) with the instructor evaluation for certain dimensions (such as Dimensions 1 and 2 in the drawing) than those who did not use bubbles. The reason is likely because students were less exposed to rubrics of drawings and assemblies than they were to rubrics of part models. Under these circumstances, the brief explanation provided by the bubbles may not have been sufficient to understand the concepts. Based on Orsmond et al. [1996], it is presumed that the concepts of drawing and assembly quality are still so alien that

many students may be unable to recognize them by solely using the explanations from the bubbles. Again, more illustrative anchors may still be required.

Table 4.9. Time (in seconds) each detailed explanation was displayed for the modeling rubric in Experiment 2.

DIMENSION 1	**No/Never**	**Almost never**	**Sometimes**	**Almost always**	**Yes/Always**	**Total**
Max time (seconds)	7.015	5.112	6.418	17.01	15.12	**38.1**
Average time (seconds)	1.661	0.901	1.287	1.74	2.713	**4.573**
Number of users of bubbles	14	14	16	19	26	**35**
DIMENSION 2						
Max time (seconds)	2.991	1.071	1.488	3.645	6.958	**6.958**
Average time (seconds)	0.842	0.551	0.46	0.82	0.955	**1.609**
Number of users of bubbles	10	12	15	16	29	**39**
DIMENSION 3						
Max time (seconds)	1.652	2.104	7.136	15.47	18.7	**42.79**
Average time (seconds)	0.525	0.668	1.297	1.224	2.15	**3.497**
Number of users of bubbles	15	5	18	27	32	**39**
DIMENSION 4						
Max time (seconds)	0.699	2.468	1.84	2.799	9.443	**9.891**
Average time (seconds)	0.239	0.588	0.505	0.683	1.637	**2.114**
Number of users of bubbles	10	14	19	24	28	**39**
DIMENSION 5						
Max time (seconds)	7.416	5.092	13.81	1.822	3.679	**15.57**
Average time (seconds)	1.384	0.91	2.541	0.529	0.836	**2.292**
Number of users of bubbles	9	14	12	23	23	**38**
DIMENSION 6						
Max time (seconds)	3.245	4.93	12.02	19.35	10.32	**49.86**
Average time (seconds)	0.738	0.843	1.54	1.513	1.204	**2.906**
Number of users of bubbles	9	12	17	31	34	**45**

It can be confirmed that the metadata captured through the rubrics platform provides valuable metrics (such as the interaction between the user and the anchors). The hypotheses that was postulated as a result of the analysis of metadata could be explored in future experiments, and if true, further work will still be required to better correlate such complex information in order to obtain valuable guidelines to improve the anchors used to recognize quality criteria.

11.3 Experiment 3

The goal of the third experiment was to test the approach to CAD training by combining suitable tutorials and materials during instruction with a set of adaptable rubrics for feedback and evaluation.

The sample was a multi-disciplinary group of senior engineering students at a US university. Students were divided into two groups: experimental (EG) and control (CG). Both groups were provided with a complete set of lecture notes, which included detailed explanations of the different dimensions of CAD quality (referred to as theoretical background). The lecture notes of the experimental group also included detailed explanations and examples of how to apply rubrics to self-evaluate classroom exercises (practical guidelines). This information was intentionally removed from the lecture notes that were given to the control group.

Participants were asked to create 3D solid models of two parts (see Figure 4.16). As an extra requirement for the second parts, participants were explicitly told to make their models flexible and reusable, allowing a series of design changes to be performed successfully and efficiently.

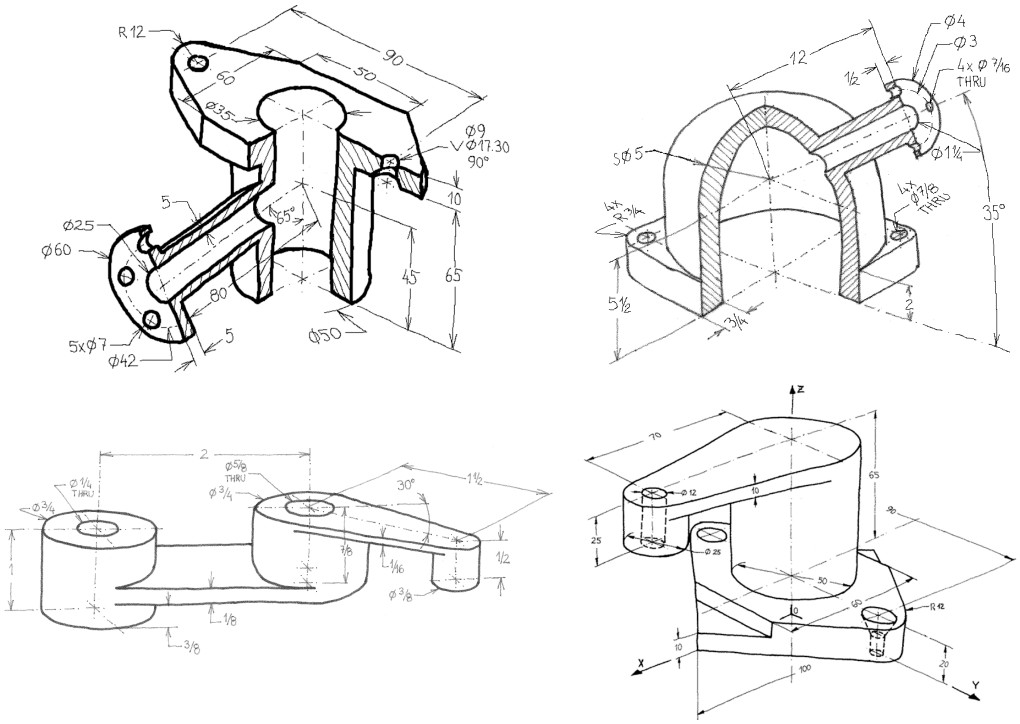


Figure 4.16. Parts used by experimental group (left) and control group (right) in Experiment 3.

All students were asked to self-evaluate their work using spreadsheet-based rubrics. After submitting the first self-evaluation, they were provided with a "solution": step-by-step instructions of an efficient strategy (defined by the research team) to model the parts, which also included detailed explanations of how to evaluate the quality of the models. With this information available, they were asked to re-evaluate their original models. The initial sample size of the study was fifty students in the experimental group and forty-nine in the control group, but only twenty-nine students in the experimental group and twenty-six in the control group completed all the required tasks.

Obviously, all the known inconveniences of spreadsheet forms appeared. As an example, 8.8% of the rubric forms (29/330) were returned unlocked and evaluated simultaneously at various levels. This problem resulted in inconsistencies between the main levels and their corresponding sub-levels. Another remarkable inconvenience was that extracting the information from the forms was an arduous, time consuming, and error-prone task. The only metadata that could be obtained from those forms was that 7 of 330 rubrics were assessed without showing the score (hence, the usefulness of such feedback was demonstrated). Finally, by analyzing their replies, it could also be determined that 53% of the rubrics always used the third level (175/330), while 8% of the assessments only used the first level (26/330). The remaining students used Level 1 for some dimensions and Levels 2 or 3 for others. This fact demonstrates that the rhythm of quality concept assimilation varies for each person and validates the vision of adaptable rubrics.

Differences were searched in the understanding of quality criteria by comparing the level of agreement between inter-rater assessments. Results in Table 4.10 reflect that inter-rater agreement between student self-evaluations of the control group (CG) and the ideal (instructor assessment) improved, as the percentage of agreement was slightly higher (9%) after the students were provided with a solution. However, there was not a homogeneous improvement for all six dimensions. The inter-rater agreement did not improve after the solution was delivered for Part 2, although it is to be noted that the levels of assessment for Part 2 before the solution was revealed were already similar to those obtained for Part 1

after the solution (62%). This fact suggests that the beneficial effect of practical guidelines exists, although it may be limited.

Table 4.10. Percentages of agreement in the control group (threshold difference of 0.25).

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
Self to Instructor before delivering the solution (part 1)	73	58	69	46	46	27	53
Self to Instructor after delivering the solution(part 1)	96	73	69	65	38	31	62
Self to Instructor before delivering the solution (part 2)	92	73	62	54	62	31	62
Self to Instructor after delivering the solution(part 2)	92	73	69	58	46	27	61

Table 4.11 reflects that improvement is similar for the experimental group (from 45% to 55% of agreement), thus bringing into question the advantage of previous exposure to similar practical guidelines before the test. Additionally, Table 4.11 shows a slight decrease (75% to 70%) in the percentages of agreement for the experimental group after the solution is delivered. So, the particular practical guidelines contained in this specific lecture note were insufficient to increase student background on practical guidelines. In fact, it even caused confusion among some of them. Apparently, the lecture notes provided were not yet fully suitable.

A weak conclusion to be made is that practical guidelines appear to have limited beneficial effect. Future studies should determine if the improvement in the agreement is due to theoretical background or to practical guidelines. Another matter of investigation would be examination of each dimension separately, in order to ascertain the sequence and balance between dimensions that produces optimal results.

Table 4.11. Percentages of agreement in the experimental group (threshold difference of 0.25).

% of agreement	**Dim 1**	**Dim 2**	**Dim 3**	**Dim 4**	**Dim 5**	**Dim 6**	**Average**
Self to Instructor before delivering the solution (part 1)	62	59	52	31	45	24	45
Self to Instructor after delivering the solution(part 1)	86	59	59	34	62	28	55
Self to Instructor before delivering the solution (part 2)	97	69	86	55	90	55	75
Self to Instructor after delivering the solution(part 2)	97	69	72	48	76	55	70

Peer evaluations were also conducted, with participants asked to evaluate the work of a random student (Table 4.12). A second instructor, not previously exposed to rubrics, was also asked to assess the students (Table 4.13). Results show that the behavior of both (peer and auxiliary instructor) were similar to those of the students, and different from those of the experienced instructor. This result suggests that previous exposure to rubrics creates a significant impact: students unexposed to rubrics are unable to recognize sophisticated quality concepts beyond an introductory level. Hence, this background should be mandatory to in order to obtain accurate assessments. The problem of

minimizing intra-rater differences and finding a reliable ideal assessment remains open for future studies.

Table 4.12. Percentages of agreement between peer students and instructor (threshold difference of 0.25).

% of agreement	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Avg.
Peer to main Instructor (CG part 1)	96	65	46	50	50	23	55
Peer to main Instructor (CG part 2)	100	85	62	54	65	19	64
Peer to main Instructor (EG part 1)	90	59	66	31	62	34	57
Peer to main Instructor (EG part 2)	97	76	90	48	83	69	77

Table 4.13. Percentages of agreement between instructors (threshold difference of 0.25).

% of agreement	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Avg.
Auxiliary to main Instructor (CG part 1)	100	81	50	58	8	15	52
Auxiliary to main Instructor (CG part 2)	100	73	38	85	50	31	63
Auxiliary to main Instructor (EG part 1)	100	72	69	10	45	21	53
Auxiliary to main Instructor (EG part 2)	97	79	90	66	52	62	74

To determine if student behavior was identical for assemblies and drawings, a second test was administered, where students were required to assemble a virtual device (all parts were provided) as shown in Figure 4.17. All evaluations were conducted using the Annota platform.

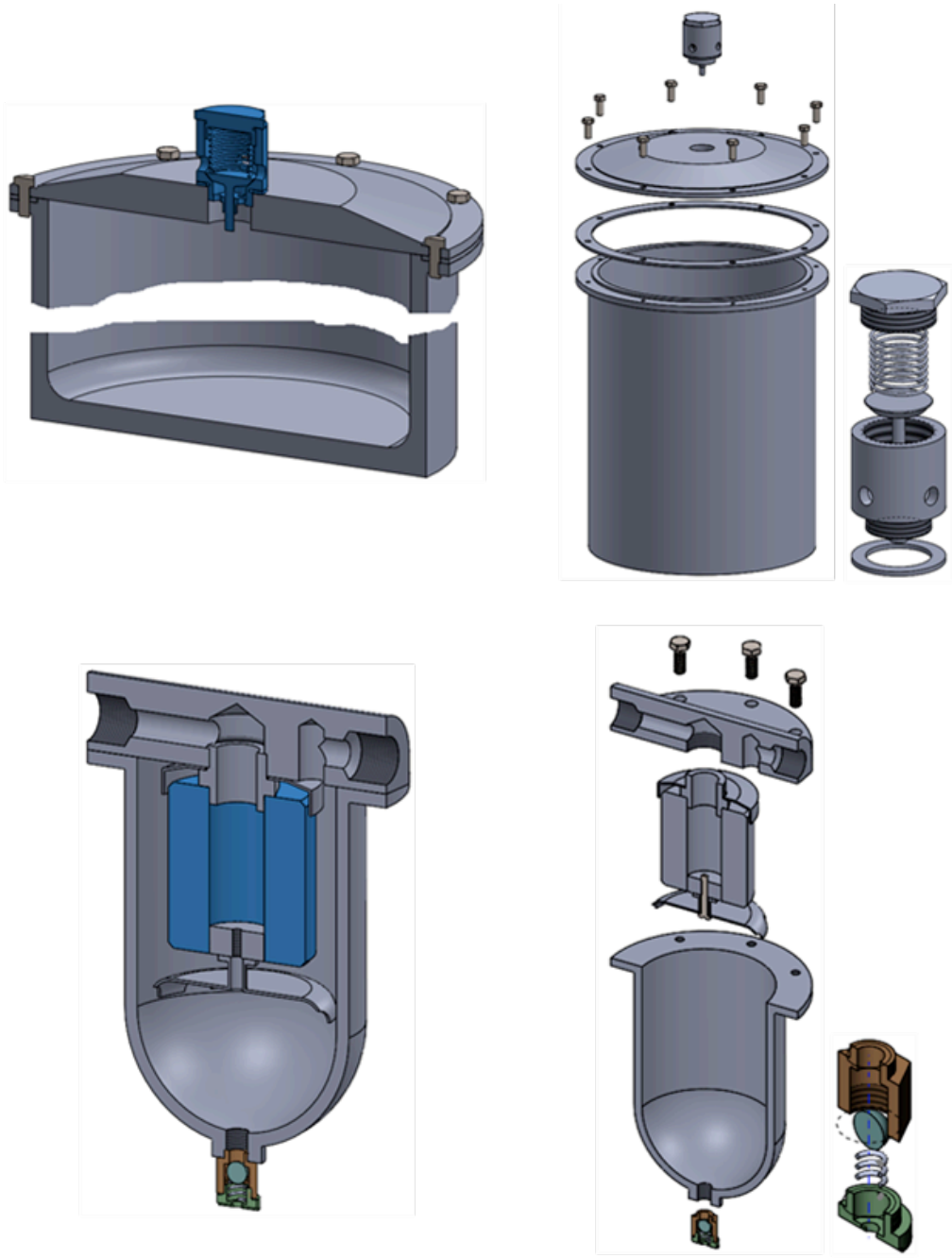


Figure 4.17. Assemblies used by the experimental group (pressure tank, in top row) and control group (air filter, in bottom row) in Experiment 3.

In the control group (CG), a subset of thirty-three students assembled the devices, self-evaluated their task (both before and after reading the solution), and were peer evaluated by another student. A threshold difference of 0.25 did not detect the actual small differences between the self-evaluations and the ideal assessment (Table 4.14). Probable causes for this increment in the correlation between students and the instructor are that all students had been previously exposed to the rubrics in the parts experiments, and they fine-tuned their work by using sub-criteria of Levels 2 and 3. Thus, only a narrow difference (less than 0.05 between the student and the ideal) was labeled "agreed" and was used to calculate the percentages of agreement in Table 4.15. The Experimental Group (EG) included a valid subset of thirty-six students and their inter-rater agreements versus two "ideals" (the peer students and the instructor) are shown in Table 4.16.

Table 4.14. Percentages of agreement in the control group (threshold difference of 0.25).

% of agreement	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Average
Self to Peer before delivering the solution	94	91	82	79	70	73	81
Self to Peer after delivering the solution	94	94	79	73	70	79	81
Self to Instructor before delivering the solution	85	82	39	67	30	36	57
Self to Instructor after delivering the solution	82	76	33	61	30	36	53

Table 4.15. Percentages of agreement in the control group (threshold difference of 0.05).

% of agreement	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Average
Self to Peer before delivering the solution	76	70	27	24	15	24	39
Self to Peer after delivering the solution	79	67	36	27	18	36	44
Self to Instructor before delivering the solution	70	36	3	21	12	9	25
Self to Instructor after delivering the solution	73	42	9	30	12	12	30

Table 4.16. Percentages of agreement in the experimental group (threshold difference of 0.05).

% of agreement	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Average
Self to Peer before delivering the solution	75	69	25	36	36	19	44
Self to Peer after delivering the solution	72	69	19	36	19	31	41
Self to Instructor before delivering the solution	89	69	3	44	31	6	40
Self to Instructor after delivering the solution	83	67	3	47	28	8	39

Table 4.15 illustrates that inter-rater agreement between self-evaluations of the students of the control group (CG) and both ideals differ in absolute value, as the instructor was stricter (less "optimistic") than the peer evaluators. It can also be noted that the percentage of agreement was slightly higher (5% in both cases) after the students were provided with a solution.

Since students in the control group had not been previously exposed to practical guidelines about quality concepts for CAD assemblies, this result would validate the hypothesis that such lecture notes are useful for students to better comprehend quality concepts. In other words: theoretical background about quality is insufficient while practical guidelines provides a positive impact.

However, Table 4.16 shows a slight decrease in the percentages of agreement for the experimental group after the solution was delivered. This fact suggests that perhaps the particular practical guidelines contained in this specific lecture note were inadequate to increase their previous background on practical guidelines. In fact, it even caused confusion for some students. Hence, it can be inferred, as was also the case in the second experiment, that the concepts of drawing and assembly quality are still so unfamiliar that many students may be unable to recognize them beyond an introductory level. It can also be conjectured that the lecture notes provided were not yet fully appropriate. In both cases, additional work is needed to provide better explanations in the future.

In summation, a weak conclusion can be made that using appropriate materials, in combination with the rubrics system during instruction, improves inter-rater agreement between students and the instructor, while allowing them to better understand quality concepts in order to self and peer evaluate CAD tasks reliably. Better detailed explanations and examples on how to apply rubrics to self-evaluate are needed to ensure that students fully understand the most subtle and sometimes contradictory dimensions of CAD quality. Optimistically, the e-rubric platform could be helpful to verify the

suitability of such future materials, particularly if the metadata that the platform provides (detection of which kind of material better reduces the inter-rater agreement) can be used.

12. Conclusions and Future Developments

In this chapter, a discussion of the design, implementation, and verification of a CAA e-rubrics system aimed at formative e-assessment [Pachler et al., 2010] has been presented. It has been found valid for formative assessment on MCAD quality. The system outputs metadata, which is helpful to analyze the evaluation process. It is also adaptable, which allows users to dynamically access multiple levels of detail for each quality criterion. In this implementation, the level of detail is controlled by buttons that allow users to reveal or conceal details upon request.

The platform lays the foundation for the future evolution of adaptable rubrics, which can easily become adaptive by establishing a set of interconnected tasks (linked to different rubrics) to automatically activate subsequent tasks that can better assist the student in the next modeling stages. This process is performed after the previous stage has been self-evaluated and the platform has analyzed the performance level attained by the student. Future possibilities could include using this system to merge CAD quality-based notes and materials (which have already been successfully tested separately) with the self-evaluation rubrics to develop on-line CAD courses, and also expanding the implementation so the system can be used indistinctly as a stand-alone tool, or integrated in the teaching/learning workflow as a plug-in for a Learning Management System (LMS) [Nakahara et al., 2014]; including online formative assessment [Gikandi et al., 2011], as it has been reported that formative feedback can make significant differences in learning outcomes in online courses [Lawton et al., 2012].

The experiments described above were performed by a research team of which I was a member. My specific responsibilities included assisting with draft versions of the rubrics, translation of the introductory rubric description provided to the students (developed by Pedro Company), liaising with the course instructor to facilitate data collection, assessing student models, and preliminary statistical analysis.

The next chapter describes the evolution of assembly rubrics, most notably hypothesizing that they need to be introduced bottom-up instead of top-down. The experiment relies on the Annota software, and this experiment is expounded and statistically validated.

Chapter 5
The Implementation of Assembly Rubrics

1. Introduction

This chapter describes and details the approach developed to convey quality-oriented modeling strategies, by use of rubrics, to beginning CAD learners undergoing instruction in the creation of assemblies. Once the parts rubrics had been created, tested, and statistically validated, the next logical step was to create rubrics that would assess assemblies. In order to remain consistent with previous research performed by the team, the parts rubrics dimensions descriptions were redefined and reused in order to be applicable to assemblies.

2. Assembly Rubric Development

For simplicity, an effort was expended to retain significant consistency between the parts and newly created assembly rubrics, but some variation exists between the two. Most notably, the descriptors and achievement levels for each criterion were reformulated to be more conducive to assemblies.

2.1 Descriptors and Achievement Levels

Descriptors can be understood to be statements that communicate the desired state of each assessed aspect. Descriptors are defined by three primary characteristics:

- They must correlate with a teachable result.
- They must correlate with an unbiased and easily mesurable result.
- They must not include other implicit descriptors.

The use of ambiguous descriptors prevents homogeneous evaluation, while explicit descriptors are required for proper evaluation to occur.

Achievement levels reflect the amount of conformity for each assessed aspect. Ideally, simple cases can be dichotomously determined (ex. specific knowledge is demonstrated or absent), but it should be feasible to assess the level of compliance through a series of tiers that discretize a continuum. Likert

Scales are useful, and commonly used, for this purpose. Achievement levels can be defined by two characteristics:

- They must use the same terms as the corresponding criteria.
- The scale should be consistent throughout all achievement levels.

All achievement levels are required to be organized in a manner such that they follow the same order throughout the rubric, either in increasing or decreasing order. These levels should be consistently described (using identical terms in the criterion), but must also be differentiated using appropriate qualifiers for each attribute. According to Rohrmann et al. [2007], qualifiers can be described by:

- Frequency (never, rarely, sometimes, often, always, etc.).
- Intensity (not at all, slightly, moderately, considerably, extremely, etc.).
- Probability (certainly not, unlikely, likely, certainly, etc.).

Figure 5.1 illustrates a sample of the assembly rubric as utilized in the experiments, showing the intensity qualifiers that describe the performance levels for validity.

CRITERION		WEIGHT	PERFORMANCE LEVEL				
#	Description	%	No/Never	Almost Never/Rarely	Sometimes	Almost always/Mostly	Yes/Always
1	**The assembly is valid (stop evaluating if not valid)**	0.00	There is no file of the assembly or linked files, or the assembly is empty	The file of the assembly, or their linked files, require intense file management before they can be located and used	The file of the assembly , or their linked files, require moderate file management before they can be located and used	The file of the assembly , and their linked files, can be located and used, but require some minor file management	The file of the assembly, and their linked files,are easily located and used

Figure 5.1. Sample assembly rubric illustrating performance levels.

2.2 Rubric Scoring

Rubrics necessarily generate scores, because scoring is consubstantial to the rubric. Scores are the essential output of the rubric (their reason for being used at all). Defining the scoring process (ex. using formulas) is required to provide the aggregated score from the achievement levels. In the research team's view, the scoring process can be improved by three characteristics:

- Dichotomous criterion is defined as when only two evaluations are reasonable: fail/pass. Ideally, the more dichotomous the scoring, the more unbiased the measurements will be, especially in situations with mulitple assessors. Dichotomous criteria also provide more opportunities for automating the scoring process.
- Evaluation criteria can have varying levels of importance (ex. different weights for each criterion), which must be made explicit in the rubric.
- Go/No Go criteria (when a failure in one criterion is so critical that it prevents analyzing other aspects of the subject's performance), may be used, but they must be explicitly identified, and included as such, in the descriptor. Go/No Go criterion can also include a threshold parameter (Ex. After ten errors, the assigned grade becomes zero, regardless of satisfying other rubric criteria.)

3. Rubric Dimensions

For the proposed assembly rubric, a classification system was developed with the following CAD modeling quality dimensions:

- An assembly is valid if it can be located, opened, and can be used with all parts accessible (established as a go/no go criterion).
- An assembly is complete if it contains all and only the necessary components, they are correctly placed, and are free of unwarranted interferences.
- An assembly is consistent if the base part is correctly assigned, valid movement is allowed, while invalid movement is prevented.
- An assembly is concise if it is free of repetitive mating conditions, uses replication operations when germane, and relationships are free of unnecessary dependencies.
- An assembly is clear if all parts and mates are labeled and organized, and compatible mates are used.
- An assembly conveys design intent if the assembly tree (history tree) replicates the assembly process, sub-assemblies (if appropriate) have been utilized, and mating features have been used to mimic the actual assembly.

Detailed discussion of each rubric dimension follows, including recommendations for best practices. Detailed educational presentations of these concepts, developed by Pedro Company, are included in the appendix. These presentations were utilized in the curricular activities prior to both experiments (described in Section 5).

3.1 Validity

The original intent was to identify validity as a "go/no go" switch, so that the assembly would fail assessment if all linked files could not be located or used. In practice, achievement levels were used to score validity, while the total score was influenced by the validity score. In this way, catastrophic validity failures result in a no go, while moderate validity failures reduce the final score, but do not prevent assessing the other rubric dimensions. This "soft" go/no go is a recommended academic scoring alternative necessary to highlight critical failures, while avoiding unnecessary punitive student exam scores (so that maximum partial credit could be awarded). Of course, industry use of these rubrics may not benefit from such allowance. An assembly is considered valid if it can be retrieved, safely used, and all references are linked.

A file is easy to locate if consistent saving practices and file naming conventions are used. Verification could include ensuring that the file contains the labeled assembly and that each part file describes its contents. A file that can be successfully accessed should open in a neutral state (without operations in progress) and files should not be manipulated while in use.

As assembly can be safely used only if it is compatible with the CAD application of the receiver (even the software version). Items to consider include whether the file is in "read only" mode or if an exported file is in a compatible format. If an assembly contains errors, the user should troubleshoot or revert to an earlier version of the file that contains no errors.

Linked files in assemblies are also required to be located and opened. Access to these parts is critical, otherwise the proper assembly will fail to be accessed. Good practices dictate that not only should these files be easily located, but that they should automatically open without searching, and that the assembly should not require rebuilding. Ideally, all assembly files should be placed in the same

folder, or "Pack and Go" (in the SolidWorks terminology) should be used so that all files, including standard library parts, will be locally available.

3.2 Completeness

An assembly can be considered complete if it includes all required components (part and sub-assembly files), uses standard component files (library) when appropriate, and all components are correctly placed.

A complete assembly must contain all necessary components. Good practices include visual inspection of the history tree to verify that all required components are provided (including multiple copies of an identical part). Using different colors for each component is a good strategy to detect the presence of all required files. Colors can be utilized to provide contrast between components (which is the best choice for inspection purposes) or to provide realistic material appearance (only recommended for rendering).

Standard components (i.e. fasteners, bearings, etc.) should be used in order to save time and effort. It is significantly easier to use the software to provide a fastener for a hole than for the user to create it from scratch. Problems occur however, when sharing assemblies with other users who have different library settings or installation. A simple solution could be individually saving each standard component as a separate part file, although if the hole is updated, additional effort is required to create another fastener. Also, in such cases, the fastener would not automatically update should the hole be altered.

All parts should be correctly placed in the assembly framework. Use of views, display styles, sections, and transparency settings are helpful to inspect whether each component is in the correct location. Some software even provide various tools which check for interference between components. Of course, interference detection needs to be personally verified (using acquired engineering experience), as some forms of interference may be required for design purposes (such as between simplified male and female threads).

3.3 Consistency

In the product design process, proper and reliable analysis can only be obtained using consistent models. Assembly models are the principal view of the digital representation of these products. Secondary views can be used for mock-up analysis and manufacturing, but the primary view must be consistent for this situation to be beneficial.

An assembly must be upright, centered, and symmetricallly placed in order for it to interact with the specific environment during analysis. These conditions are also important when a sub-assembly must be placed and function within a larger assembly. Since the base part (or parts) behaves as a physical anchor for the assembly, and is frequently fixed, it must be linked to the global reference system.

All components should be suitably mated to ensure proper placement, with attention given to removing only the degrees of freedom necessary to mimic actual mechanisms. The assembly must allow for valid motion, while simultaneously preventing invalid motion. Both requirements must be satisfied in order for proper analysis to occur (kinematic, structural, frequency, thermal, etc.).

3.4 Conciseness

Concise assemblies do not contain repetitive or fragmented mating conditions. Mates are considered repetitive if they re-constrain the same degree of freedom. As an example, if a cylinder is presently concentric with a hole, it is redundant to add a coaxial mate between the features. Fragmented mates

should also be avoided, as multiple simple mates are less efficient than one comprehensive one. As in most modeling situations, a balance is required in order to select the most advantageous mating scenario. Furthermore, using unnecessary mates is incorrect, but fragmenting complex mates may improve clarity while allowing for easier design exploration when configuring mechanisms.

Highly semantic assembly operations provide context between user intentions and the contents of computational assemblies. These operations provide important design information that assist those that did not create the model to analyze and manipulate the assemblies. Good practices include using pattern operations, when appropriate, to insert and link components that are arranged at regular intervals (linear, circular, and symmetry).

Construction of long chains of mating relationships between components is discouraged, as unforeseen relationships may result, while also increasing calculation times as the software becomes more prone to round off errors. It is preferable to use a small subset of base parts and relate the remaining components directly. Indirect mating is not desirable, as the mating procedure becomes more difficult and prevents editing mates when rearranging the assembly during design exploration.

3.5 Clarity

Clarity is required of assemblies because they are design documents that are shared between stakeholders throughout the design and manufacturing process. For communication to occur, the document (assembly) must be easily understood (preferably at the first viewing). This communication is facilitated if the mating operations are intelligently labeled to indicate their function and grouped to emphasize their relationships. As a rule, the most compatible and standard mating operations are always desired. While mating operations are automatically labeled in the history tree (regardless of the software), the system only provides information about how the mates were created, not their function, which is significantly more important in the communication process. It is recommeded that all mates be re-labeled to emphasize just what exactly is linked, not the type of link implemented.

Mating operations should be grouped according to the design criteria needed to increase communication. This process could be accomplished by grouping by parts or degrees of freedom. While an optimum grouping procedure does not exist, it is more important to avoid clearly erroneous solutions.

When deciding which mate to use, always select the simplest and most compatible choice. As an example, use a coincident mate rather than using zero distance. If possible, use high-level mates if they are standard or common. Agreements are required, as some mates may reduce the portability of the assembly when shared among a design team.

3.6 Design Intent

Design intent is the most advanced (and most difficult) rubric dimension to assess, as it relies not only on the modeling procedure, but also on an intricate understanding of the design's function. Many design methodologies use assembly models to investigate design behavior. Assemblly models convey design intent when they convey information that is useful for analyses. Four different aspects can be analyzed:

- Assembly planning, when assembly sequence is paramount.
- Assembly Process Design (APD), when funcionalties are examined.
- Design for Assembly (DFA), when affordances used to assemble and disassemble are analyzed.
- Varieties, which considers product families rather than isolated products.

Assembly planning is the process of creating a set of instructions used to mechanically assemble a product from a group of components. This assembly algorithm specifies the sequence of assembly, disassembly, and repair procedures. Sequencing is the most vital concept of assembly planning and must be reflected in the software's history tree in order to replicate the process. The history tree must be inspected to ensure that the assembly sequence accurately reflects authentic assembly procedures. Best practices include sequencing the assembly components from main to auxiliary elements and that the disassembly process should be inferred by reversing the history tree. A realistic assembly sequence could result in unrealistic mates, so agreement between actual sequencing and reasonable mate linking is imperative.

Component modules are useful to convey proper functionality, especially when they perform with minimal interaction with other components or sub-assemblies. These modules should be adequately labeled to reflect their purpose. Sub-assemblies can then be utilized to encapsulate these modules. Best practices suggest that mates within sub-assemblies provide for proper motion and therefore should be manipulated so they will behave as flexible mechanisms.

Design for Assembly (DFA) is a methodology in which components contain affordances (features on parts used to grasp, move, orient, and insert) which simplify the assembly process. Components which possess these mating features should exploit these affordances in the assembly process (e.g. a flap that fits into a groove).

Assembly Process Design (APD) focuses on product functionality. Process-based approaches increase the flexibility of industrial assemblies, but cannot be simulated with assembly modeling software. Product-based approaches standardize a majority of components, while providing variations for the remaining parts. Virtual components in the assembly model should be as easy to replace as the actual parts in the real world assembly. In order to meet this requirement, the indepedence of replaceable parts should be increased.

3.7 Summary

To summarize, the quality dimensions, including sub-dimensions follow:

1. The assembly is valid (this is a soft go/no go criterion, which will multiply the overall score obtained by the rest of the rubric).
    1.1. The file of the assembly can be located and opens in a neutral state.
        1.1.a. The file of the assembly has the expected contents (and name) and is in the expected place (folder or website).
        1.1.b. The file of the assembly can be re-opened after closing the current session (even on a different computer).
        1.1.c. The file of the assembly opens in a neutral state (i.e. no operations in progress).
    1.2. The assembly can be used.
        1.2.a. The assembly is compatible with the CAD program (and software version) used by the receiver.
        1.2.b. The assembly is free of error messages.
    1.3. All components (parts and sub-assemblies) linked to the assembly may be accessed, even when libraries are not available, or when software compatability issues exist between versions.

1.3.a. All parts linked to the assembly can be accessed.

1.3.b. All sub-assemblies linked to the assembly can be accessed.

1.3.c. All library components linked to the assembly can be accessed.

2. The assembly is complete.

2.1. The assembly includes all and only the necessary components (parts, sub-assemblies, and library components).

2.1.a. The assembly includes all the components and their corresponding copies.

2.1.b. The assembly is free from surplus and alien components.

2.2. Standard library components are included when required, which are suitably instantiated from the library.

2.2.a. Standard library components are used when required.

2.2.b. Standard library components are suitably instantiated from the library.

2.3. Components (parts, sub-assemblies, and library components) are correctly placed.

2.3.a. Relative locations among components match their functional positions.

2.3.b. Components are free of unwanted interferences.

3. The assemblly is consistent.

3.1. The base component is correctly assigned and is well linked to the global reference system.

3.1.a. The component selected as the base is suitable, as it acts as a support or a container and is preferably a fixed part (particularly if the assembly is a mechanism).

3.1.b. The base component is correctly linked to the global reference system, as it is centered and maximizes symmetry.

3.2. Assembly mate conditions allow valid movements while preventing undesired movements.

3.2.a. Assembly mates prevent invalid movement.

3.2.b. Assembly mates allow valid movement.

4. The assembly is concise.

4.1. The assembly is free from repetitive or fragmented mating conditions.

4.2. Replication operations (translate-and-repeat, rotate-and-repeat, and symmetry) are used whenever possible.

4.2.a. 3D patterns operations (translate-and-repeat, rotate-and-repeat) are used whenever possible.

4.2.b. Symmetry (if it exists) is used to define the assembly.

4.3. The parent/child relationships in the assembly tree are free of unnecessary dependencies.

5. The assembly is clear.

5.1. All components and mates are properly labeled and organized in folders.

5.1.a. Components are labeled and grouped to emphasize their function, instead of how they were defined.

5.1.b. Mates are labeled to emphasize their function.

5.1.c. Related mates are grouped to emphasize parent/child relationships.

5.2. The assembly uses compatible and standard mates.

5.2.a. The most compatible mates are always used.

5.2.b. The most standard mates are always used.

6. The assembly conveys design intent.

6.1. The assembly tree replicates the assembly/disassembly process.

6.1.a. The assembly sequence proceeds from main to auxiliary elements.

6.1.b. The assembly sequence reflects a realistic mounting sequence.

6.2. Sub-assemblies have been properly identified and efficiently used.

6.2.a. Sub-assemblies encapsulate clearlly perceived functions.

6.2.b. The mates of sub-assemblies provide for proper motion and have been made flexible.

6.3. Mating features provided as affordances to ease assembly, if any, are mostly used for mating.

6.3.a. Mating features provided to grasp, move, orient, and insert the part, if any, have been identifed.

6.3.b. Mating features provided to grasp, move, orient, and insert the part, if any, are mostly used for mating.

4. Rubric Rating Scale

To accommodate their varying levels of importance, the dimensions were rated as follows:

- Valid: 0% (soft go/no go criterion that multiplies the overall score obtained using the remaining rubric dimenisons).
- Complete: 20%
- Consistent: 30%
- Concise: 20%
- Clear: 15%
- Design Intent: 15%.

Rohrmann [2007] states that category scaling enhances the usability of assessment instruments and that well-defined qualifiers provide for unbiased judgments. With those concepts in mind, performance levels were defined as:

- No/Never
- Almost Never/Rarely
- Sometimes
- Almost Always/Mostly
- Yes/Always.

While objective scoring is difficult to obtain, especially for those who are self-assessing, these performance level categories provide unambiguous scales to properly rate model quality. When assessing student performance, a preferred strategy involves moderate leniency when awarding scores, in order to build confidence in beginning CAD users. Instead of viewing a specific, small error as important enough to prevent awarding a top score, a proper assessment perspective could involve viewing individual instances of small faults as not important enough to prevent awarding a maximum rating.

5. Experiments

Two experiments were conducted (mid-term and final exam) assessing student understanding of assembly rubrics using Annota e-Rubrics. Annota e-Rubrics were discussed in detail in Chapter 4 (Section 4.10.2). The experiments demonstrated stronger agreement between instructors than either instructor with the students, for all dimensions. Agreement between instructors and students was obtained for the dimensions of validity, completeness, and clarity, but weak agreement exists for consistency, conciseness, and design intent.

5.1 Experiment 1

Undergraduate students (beginning CAD users) at a Spanish university were introduced to prototype assembly rubrics, having been exposed to parts rubrics earlier in the semester. Detailed explanations of the assembly rubric dimensions (developed by Pedro Company and included in the appendix) were discussed and provided to the students prior to their examinations. This introductory material included thorough descriptions of the definition and significance of the six quality dimensions, with further clarifications of the detailed criteria used to measure the degree of accomplishment of such dimensions. As stated earlier in this chapter, these quality dimensions were aligned with preceding research on parts rubrics, accomplished by the same research team.

Completion of Annota rubrics were required and considered correct if they matched the primary instructor (Instructor 1) evaluation (ideal). The primary instructor (Instructor 1) was the professor of record for the course and Instructor 2 was a faculty member at another institution, whose sole responsibility was to assess the student work.

Fifty-two students were enrolled in the class, but only fifty students sat for the exam, with only forty-six students submitting self-assessment rubrics. Students were required to assemble a fitness equipment pulley, using four non-standard parts (previously modeled) and various standard parts. The students were specifically warned on assembly sequence and also on the use of sub-assemblies. Standard parts included four hexagon socket head cap screws (ISO 4762 M3x8-8), fourteen radial ball bearings (ISO 15 RBB, size 2025), two support rims (DIN 988, size 25x35 mm), and one lock washer (DIN 6799, with 19mm groove diameter). Non-standard parts (which are shown in Figure 5.2) are an L-Bracket, Bolt, Base, and Wheel.

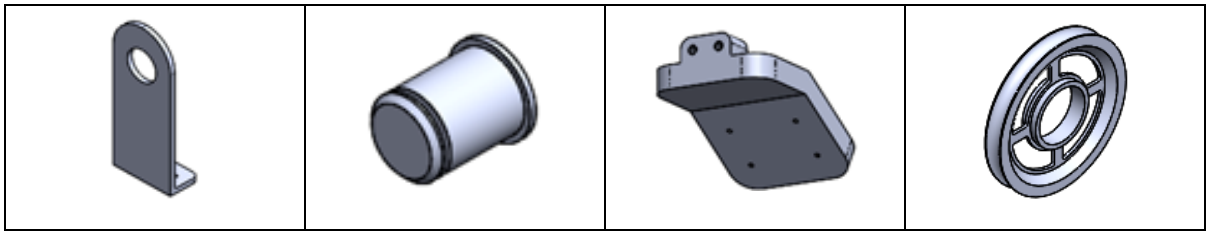


Figure 5.2. Non-standard parts used for modeling in Experiment 1. Part 1 (L-Bracket, far left), Part 2 (Bolt, center left), Part 3 (Base, center right), and Part 4 (Wheel, far right).

Two sub-assemblies were assumed: (1) Anchor Arm consisting of Base, L-Brackets, and Fixing Screws and (2) Bearing Wheel consisting of Wheel and Bearing. The expected sub-assemblies are shown in Figure 5.3, with the final assembly solution provided in Figure 5.4. The students were provided the solution after exam submittal in order to judge their performance against an ideal.

The students were informed that Dimension 1 (validity) would be a "hard" go/no go criterion, meaning that failure to submit a valid file would result in a non-passing grade for the exam. However, a "soft" go/no go criterion was enforced (with up to half-credit being awarded to avoid unnecessarily punitive scoring).

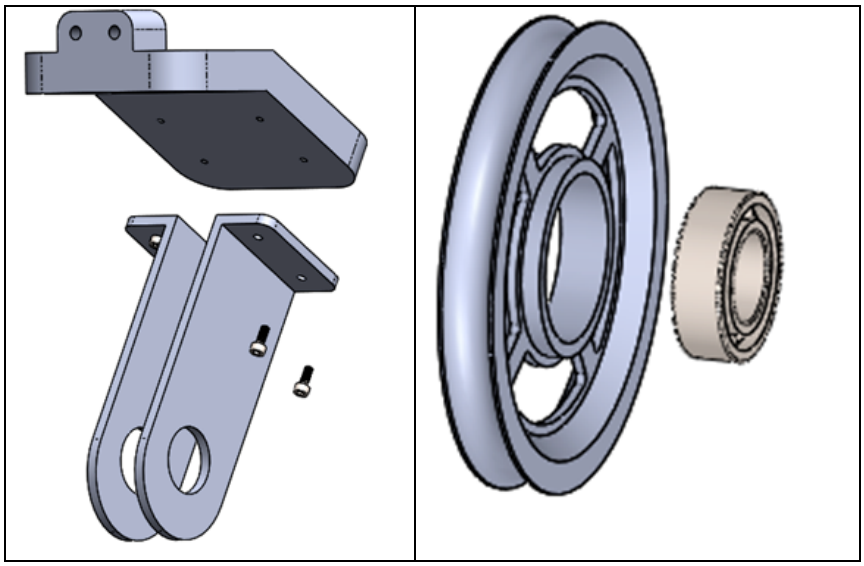


Figure 5.3. Sub-assemblies used in Experiment 1. Anchor Arm (left) and Bearing Wheel (right).

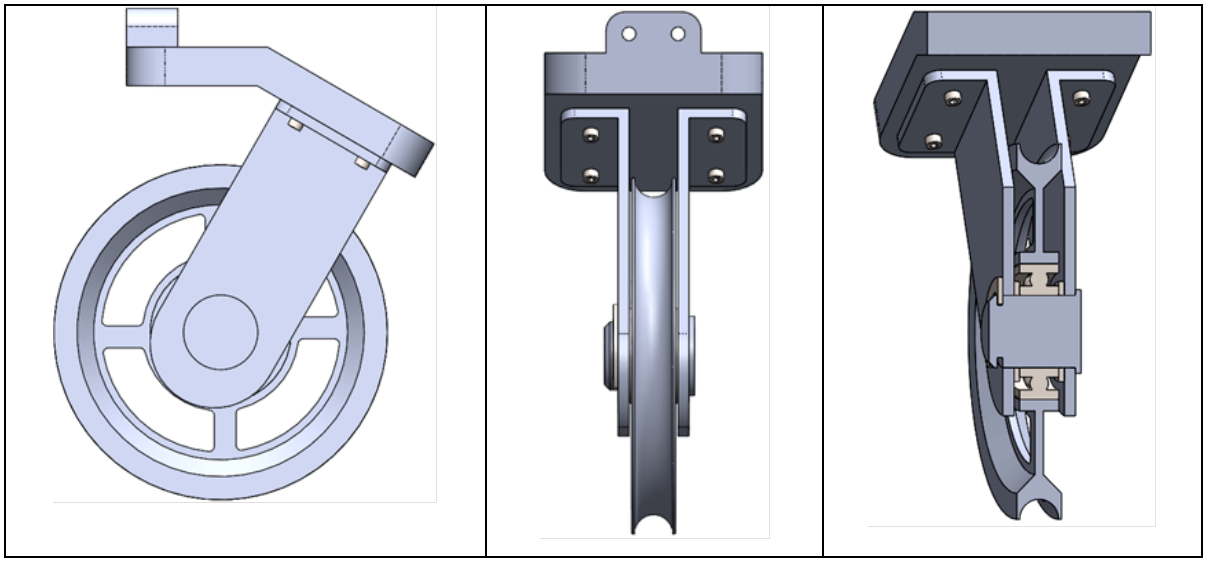


Figure 5.4. Final assembly in Experiment 1. Front View (left), Side View (center), and Sectioned Pictorial (right).

5.2 Experiment 1 Discussion

Tables 5.1 through 5.6 show the assessment for this task (for each criteria), performed by each student and instructor. Also shown are the instructor average, the difference in scores between instructors, and the difference in scores between the instructor average and each student.

Table 5.1. Validity scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1-Individual	Δ Instructor 2-Individual	Δ Instructor 1-Instructor 2	Δ Instructor Ave-Individual
1		0.50	0.50	0.50			0.00	
2								
3	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
4		1.00	1.00	1.00				
5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
6	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00
7	0.94	0.50	0.50	0.50	-0.44	-0.44	0.00	-0.44
8		1.00	1.00	1.00				
9	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
10	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
11	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
12		1.00	1.00	1.00			0.00	
13	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
14	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
15	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.89	1.00	1.00	1.00	0.11	0.11	0.00	0.11
18	1.00	0.75	1.00	0.88	-0.25	0.00	-0.25	-0.13
19	0.67	0.25	0.50	0.38	-0.42	-0.17	-0.25	-0.29
20	0.97	1.00	1.00	1.00	0.03	0.03	0.00	0.03
21	1.00	0.25	0.00	0.13	-0.75	-1.00	0.25	-0.88
22	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
23	0.79	0.50	0.50	0.50	-0.29	-0.29	0.00	-0.29
24	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
25	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
26	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
29	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
30	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
31	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
32	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
33	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
34	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
35	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
36	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
37	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
38	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
39	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
40	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
41	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
42	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
43	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
44	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
47	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
48	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
49								
50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
51	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
52	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
				AVE	-0.16	-0.15	-0.01	-0.15

Table 5.2. Completeness scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1		1.00	1.00	1.00			0.00	
2								
3	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
4		0.83	1.00	0.92			-0.17	
5	1.00	1.00	0.83	0.92	0.00	-0.17	0.17	-0.08
6	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
7	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
8		1.00	1.00	1.00			0.00	
9	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
10	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
11	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
12		0.83	0.92	0.88			-0.08	
13	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
14	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
15	0.50	1.00	0.83	0.92	0.50	0.33	0.17	0.42
16	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
17	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
18	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
19	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
20	0.96	0.92	1.00	0.96	-0.04	0.04	-0.08	0.00
21	0.25	0.00	0.00	0.00	-0.25	-0.25	0.00	-0.25
22	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
23	0.83	0.83	1.00	0.92	0.00	0.17	-0.17	0.08
24	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
25	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
26	1.00	0.75	0.83	0.79	-0.25	-0.17	-0.08	-0.21
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	0.42	0.67	0.83	0.75	0.25	0.42	-0.17	0.33
29	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
30	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
31	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
32	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
33	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
34	1.00	0.83	1.00	0.92	-0.17	0.00	-0.17	-0.08
35	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
36	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
37	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
38	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
39	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
40	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
41	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
42	1.00	0.83	0.92	0.88	-0.17	-0.08	-0.08	-0.13
43	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
44	0.92	0.58	0.83	0.71	-0.33	-0.08	-0.25	-0.21
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
47	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
48	0.75	0.92	1.00	0.96	0.17	0.25	-0.08	0.21
49								
50	0.50	0.83	1.00	0.92	0.33	0.50	-0.17	0.42
51	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
52	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
				AVE	0.11	0.13	-0.02	0.12

Table 5.3. Consistency scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1		1.00	1.00	1.00			0.00	
2								
3	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
4		1.00	1.00	1.00			0.00	
5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
6	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
7	0.63	0.50	0.50	0.50	-0.13	-0.13	0.00	-0.13
8		1.00	1.00	1.00			0.00	
9	0.50	0.50	0.63	0.56	0.00	0.13	-0.13	0.06
10	0.94	1.00	1.00	1.00	0.06	0.06	0.00	0.06
11	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
12		1.00	1.00	1.00			0.00	
13	0.75	0.00	0.13	0.06	-0.75	-0.63	-0.13	-0.69
14	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
15	0.75	1.00	0.88	0.94	0.25	0.13	0.13	0.19
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.81	1.00	1.00	1.00	0.19	0.19	0.00	0.19
18	0.94	1.00	1.00	1.00	0.06	0.06	0.00	0.06
19	0.63	1.00	0.88	0.94	0.38	0.25	0.13	0.31
20	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
21	0.25	0.00	0.00	0.00	-0.25	-0.25	0.00	-0.25
22	0.75	0.13	0.25	0.19	-0.63	-0.50	-0.13	-0.56
23	0.69	0.50	1.00	0.75	-0.19	0.31	-0.50	0.06
24	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
25	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
26	1.00	0.13	0.25	0.19	-0.88	-0.75	-0.13	-0.81
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
29	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
30	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
31	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
32	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
33	0.75	0.50	0.63	0.56	-0.25	-0.13	-0.13	-0.19
34	0.75	0.75	1.00	0.88	0.00	0.25	-0.25	0.13
35	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
36	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
37	1.00	0.88	1.00	0.94	-0.13	0.00	-0.13	-0.06
38	0.94	0.50	0.50	0.50	-0.44	-0.44	0.00	-0.44
39	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
40	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
41	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
42	1.00	0.88	1.00	0.94	-0.13	0.00	-0.13	-0.06
43	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
44	1.00	0.88	0.75	0.81	-0.13	-0.25	0.13	-0.19
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
47	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
48	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
49								
50	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
51	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
52	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
				AVE	0.01	0.04	-0.03	0.03

Table 5.4. Conciseness scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1		0.50	0.33	0.42			0.17	
2								
3	0.58	0.67	0.67	0.67	0.08	0.08	0.00	0.08
4		0.58	0.50	0.54			0.08	
5	0.75	0.50	0.42	0.46	-0.25	-0.33	0.08	-0.29
6	0.25	0.58	0.58	0.58	0.33	0.33	0.00	0.33
7	0.83	0.83	0.83	0.83	0.00	0.00	0.00	0.00
8		0.42	0.42	0.42			0.00	
9	0.50	0.67	0.67	0.67	0.17	0.17	0.00	0.17
10	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
11	0.75	0.92	1.00	0.96	0.17	0.25	-0.08	0.21
12		0.58	0.58	0.58			0.00	
13	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
14	0.50	0.58	0.58	0.58	0.08	0.08	0.00	0.08
15	0.75	0.58	0.58	0.58	-0.17	-0.17	0.00	-0.17
16	0.83	0.92	1.00	0.96	0.08	0.17	-0.08	0.13
17	0.50	0.58	0.58	0.58	0.08	0.08	0.00	0.08
18	0.67	0.58	0.58	0.58	-0.08	-0.08	0.00	-0.08
19	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00
20	0.42	0.33	0.42	0.38	-0.08	0.00	-0.08	-0.04
21	0.50	0.00	0.00	0.00	-0.50	-0.50	0.00	-0.50
22	0.75	0.67	0.67	0.67	-0.08	-0.08	0.00	-0.08
23	0.50	0.58	0.58	0.58	0.08	0.08	0.00	0.08
24	0.75	0.42	0.42	0.42	-0.33	-0.33	0.00	-0.33
25	0.83	0.83	0.92	0.88	0.00	0.08	-0.08	0.04
26	1.00	0.42	0.42	0.42	-0.58	-0.58	0.00	-0.58
27	1.00	0.92	1.00	0.96	-0.08	0.00	-0.08	-0.04
28	0.33	0.58	0.58	0.58	0.25	0.25	0.00	0.25
29	1.00	0.58	0.58	0.58	-0.42	-0.42	0.00	-0.42
30	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
31	0.42	0.67	0.67	0.67	0.25	0.25	0.00	0.25
32	0.75	0.67	0.67	0.67	-0.08	-0.08	0.00	-0.08
33	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
34	0.75	0.83	0.92	0.88	0.08	0.17	-0.08	0.13
35	0.75	0.67	0.67	0.67	-0.08	-0.08	0.00	-0.08
36	1.00	0.58	0.67	0.63	-0.42	-0.33	-0.08	-0.38
37	0.83	0.83	1.00	0.92	0.00	0.17	-0.17	0.08
38	0.63	0.67	0.67	0.67	0.04	0.04	0.00	0.04
39	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
40	0.75	0.83	0.92	0.88	0.08	0.17	-0.08	0.13
41	0.67	0.58	0.58	0.58	-0.08	-0.08	0.00	-0.08
42	1.00	0.33	0.50	0.42	-0.67	-0.50	-0.17	-0.58
43	0.92	0.75	0.83	0.79	-0.17	-0.08	-0.08	-0.13
44	0.83	0.67	0.67	0.67	-0.17	-0.17	0.00	-0.17
45	0.75	0.92	1.00	0.96	0.17	0.25	-0.08	0.21
46	0.75	0.58	0.67	0.63	-0.17	-0.08	-0.08	-0.13
47	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
48	1.00	0.58	0.58	0.58	-0.42	-0.42	0.00	-0.42
49								
50	0.75	0.33	0.42	0.38	-0.42	-0.33	-0.08	-0.38
51	0.42	0.83	0.92	0.88	0.42	0.50	-0.08	0.46
52	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
				AVE	-0.07	-0.04	-0.03	-0.05

Table 5.5. Clarity scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1-Individual	Δ Instructor 2-Individual	Δ Instructor 1-Instructor 2	Δ Instructor Ave-Individual
1		0.38	0.38	0.38			0.00	
2								
3	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
4		1.00	1.00	1.00			0.00	
5	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
6	0.75	0.88	1.00	0.94	0.13	0.25	-0.13	0.19
7	0.94	1.00	1.00	1.00	0.06	0.06	0.00	0.06
8		0.63	0.63	0.63			0.00	
9	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
10	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
11	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
12		0.50	0.50	0.50			0.00	
13	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00
15	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.63	0.88	0.75	0.81	0.25	0.13	0.13	0.19
18	0.58	0.75	0.88	0.81	0.17	0.29	-0.13	0.23
19	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
20	0.42	0.50	0.88	0.69	0.08	0.46	-0.38	0.27
21	0.50	0.00	0.00	0.00	-0.50	-0.50	0.00	-0.50
22	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
23	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
24	1.00	0.75	0.75	0.75	-0.25	-0.25	0.00	-0.25
25	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
26	1.00	0.88	0.88	0.88	-0.13	-0.13	0.00	-0.13
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	0.38	0.75	0.88	0.81	0.38	0.50	-0.13	0.44
29	1.00	0.63	0.75	0.69	-0.38	-0.25	-0.13	-0.31
30	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
31	0.38	0.75	0.88	0.81	0.38	0.50	-0.13	0.44
32	1.00	0.75	0.88	0.81	-0.25	-0.13	-0.13	-0.19
33	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
34	1.00	0.63	0.38	0.50	-0.38	-0.63	0.25	-0.50
35	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
36	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
37	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
38	0.83	0.75	0.88	0.81	-0.08	0.04	-0.13	-0.02
39	0.75	0.88	0.88	0.88	0.13	0.13	0.00	0.13
40	1.00	0.50	0.75	0.63	-0.50	-0.25	-0.25	-0.38
41	1.00	0.50	0.75	0.63	-0.50	-0.25	-0.25	-0.38
42	1.00	0.25	0.38	0.31	-0.75	-0.63	-0.13	-0.69
43	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
44	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
47	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
48	0.75	0.50	0.38	0.44	-0.25	-0.38	0.13	-0.31
49								
50	0.75	0.75	1.00	0.88	0.00	0.25	-0.25	0.13
51	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
52	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
				AVE	0.00	0.05	-0.04	0.03

Table 5.6. Design Intent scores for mid-term exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1		0.17	0.25	0.21			-0.08	
2								
3	0.33	0.42	0.58	0.50	0.08	0.25	-0.17	0.17
4		0.42	0.42	0.42			0.00	
5	0.50	0.25	0.33	0.29	-0.25	-0.17	-0.08	-0.21
6	0.75	0.25	0.42	0.33	-0.50	-0.33	-0.17	-0.42
7	0.79	0.42	0.58	0.50	-0.38	-0.21	-0.17	-0.29
8		0.25	0.42	0.33			-0.17	
9	0.50	0.17	0.33	0.25	-0.33	-0.17	-0.17	-0.25
10	1.00	0.67	0.75	0.71	-0.33	-0.25	-0.08	-0.29
11	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
12		0.42	0.42	0.42			0.00	
13	1.00	0.58	0.58	0.58	-0.42	-0.42	0.00	-0.42
14	0.50	0.08	0.33	0.21	-0.42	-0.17	-0.25	-0.29
15	0.50	0.33	0.33	0.33	-0.17	-0.17	0.00	-0.17
16	0.67	0.50	0.50	0.50	-0.17	-0.17	0.00	-0.17
17	0.38	0.08	0.33	0.21	-0.29	-0.04	-0.25	-0.17
18	0.67	0.25	0.33	0.29	-0.42	-0.33	-0.08	-0.38
19	0.42	0.50	0.42	0.46	0.08	0.00	0.08	0.04
20	0.21	0.33	0.33	0.33	0.13	0.13	0.00	0.13
21	0.50	0.00	0.00	0.00	-0.50	-0.50	0.00	-0.50
22	0.50	0.25	0.33	0.29	-0.25	-0.17	-0.08	-0.21
23	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
24	0.50	0.17	0.33	0.25	-0.33	-0.17	-0.17	-0.25
25	0.83	0.58	0.58	0.58	-0.25	-0.25	0.00	-0.25
26	0.75	0.25	0.42	0.33	-0.50	-0.33	-0.17	-0.42
27	1.00	0.83	0.92	0.88	-0.17	-0.08	-0.08	-0.13
28	0.58	0.33	0.42	0.38	-0.25	-0.17	-0.08	-0.21
29	1.00	0.33	0.42	0.38	-0.67	-0.58	-0.08	-0.63
30	0.75	0.42	0.50	0.46	-0.33	-0.25	-0.08	-0.29
31	0.50	0.42	0.50	0.46	-0.08	0.00	-0.08	-0.04
32	0.75	0.17	0.50	0.33	-0.58	-0.25	-0.33	-0.42
33	0.75	0.25	0.42	0.33	-0.50	-0.33	-0.17	-0.42
34	0.75	0.25	0.58	0.42	-0.50	-0.17	-0.33	-0.33
35	0.50	0.33	0.42	0.38	-0.17	-0.08	-0.08	-0.13
36	0.75	0.42	0.67	0.54	-0.33	-0.08	-0.25	-0.21
37	0.96	0.33	0.50	0.42	-0.62	-0.46	-0.17	-0.54
38	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
39	0.75	0.33	0.42	0.38	-0.42	-0.33	-0.08	-0.38
40	0.75	0.25	0.42	0.33	-0.50	-0.33	-0.17	-0.42
41	0.92	0.17	0.33	0.25	-0.75	-0.58	-0.17	-0.67
42	1.00	0.42	0.58	0.50	-0.58	-0.42	-0.17	-0.50
43	0.58	0.17	0.33	0.25	-0.42	-0.25	-0.17	-0.33
44	0.42	0.42	0.42	0.42	0.00	0.00	0.00	0.00
45	0.75	0.42	0.42	0.42	-0.33	-0.33	0.00	-0.33
46	0.75	0.17	0.33	0.25	-0.58	-0.42	-0.17	-0.50
47	1.00	0.42	0.50	0.46	-0.58	-0.50	-0.08	-0.54
48	1.00	0.42	0.50	0.46	-0.58	-0.50	-0.08	-0.54
49								
50	0.75	0.08	0.33	0.21	-0.67	-0.42	-0.25	-0.54
51	0.58	0.42	0.42	0.42	-0.17	-0.17	0.00	-0.17
52	0.50	0.42	0.42	0.42	-0.08	-0.08	0.00	-0.08
				AVE	-0.34	-0.23	-0.11	-0.28

For the dimension of Validity (Table 5.1), it can be shown that the instructors were more pessimistic than the students, but both instructors provided similar assessments. For the dimension of Completeness (Table 5.2), the instructors were more optimistic than the students, with both instructors comparably. Similar assessments were provided by students and instructors for the dimensions of Consistency (within 3%, shown in Table 5.3), Conciseness (within 5%, shown in Table 5.4), and Clarity (within 4%, shown in Table 5.5). Both instructors were strongly more pessimistic than the students for the dimension of Design Intent (Table 5.6), which is to be expected as it is a difficult concept for beginning learners to grasp. Instructor 1 was more pessimistic than Instructor 2 in this regard.

The first hypothesis to validate or reject was that assembly rubrics produce an objective accumulative assessment of students. In order to validate this hypothesis, the assesment performed by Instructor 1 was compared against the assesment made by Instructor 2 (instructor inter-rater reliability and and Pearson Correlation).

Since future improvements seem to be necessary to acquire full validity of the rubric for formative purposes, a detailed qualitative analysis was conducted to determine at what extent the designed assembly rubric is currently valid for formative purposes.

While the developed rubrics were primarily created to assess CAD model quality, the rubrics themselves can be assessed for ease of understanding and use (which is an underlying research hypothesis). If a rubric is clearly understood, each rater (instructor and student) should produce similar assessments. If there is substantial variation between raters, the reliability of scientific studies could come into question [Gwet, 2014]. The advantage is that if inter-rater reliability is high, raters can be used interchangeably [Gwet, 2014], thus reinforcing the belief that the rubrics are easily understood and applied. The requirement for rater interchangeability is paramount so that wide-spread rubric adoption can be achieved.

Table 5.7 illustrates the inter-rater reliability scores for the mid-term exam (for the student and both instructors). At first glance, it can been seen that there is greater agreement between the instructors than between instructor and students. Dimension 1 provides the most agreement and diminishes through Dimensions 2-6. Dimension 6 (design intent) provides the least agreement (between both instructors and students) and is perhaps due to its more difficult comprehension.

The inter-rater reliability between instructors was high for most dimensions, but it cannot be assumed that this agreement was related to time working together, as both instructors were separated by great distances and only corresponded electronically. This agreement can rather be explained by the use of well-designed rubrics utilized by knowledgeable instructors with detailed understanding of the process of teaching CAD.

Table 5.7. Inter-rater reliabilty scores for mid-term exam.

% of Agreement	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	50.0	36.5	27.0	17.0	23.0	7.6
Individual-Instructor 2	51.9	38.0	28.8	21.0	21.0	11.5
Instructor 1-Instructor 2	94.0	75.0	73.0	61.5	63.0	28.8

Table 5.8 displays the Pearson Correlation values for the mid-term exam (for the student and both instructors). Initially, it is revealed that very strong correlation exists between the instructors, but less so between each instructor and the students. Specifically, there exists extremely high correlation between instructors for Dimensions 1 through 5, and high correlation for Dimension 6. The slight decrease in correlation could be again, related to the more difficult concept of design intent. In examining the correlation between the instructors and students, there is moderate correlation for Dimensions 1 and 2 for Instructor 1, and between Dimensions 1, 2, 3, and 6 for Instructor 2. There is

weak correlation between instructors and students for Dimension 5 (for both instructors) and Dimension 6 for Instructor 1. The weakest correlation is for Dimension 5.

Table 5.8. Pearson Correlation values for mid-term exam.

Correlation Coefficient	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	0.59945	0.59943	0.49836	0.47598	0.36862	0.43827
Individual-Instructor 2	0.55992	0.62977	0.52457	0.50092	0.32739	0.60178
Instructor 1-Instructor 2	0.97734	0.93756	0.95118	0.97104	0.92310	0.84657

Results from the mid-term exam illustrate similar behavior in the evaluations of both instructors. Thus, it can be concluded that the designed assembly rubric is homogeneous for accumulative evaluation of CAD assembly. Results also illustrate partially similar behavior between instructor and student evaluations. Thus, it can be concluded that the designed assembly rubric has limited validity for formative self-evaluation of CAD assembly, as agreement between instructors and students was obtained for the dimensions of validity, completeness, and clarity, but weak agreement exists for consistency, conciseness, and design intent.

To shed further light on any relevant information that may have been overlooked, the research team searched for differences in the understanding of quality criteria by comparing significant differences between inter-rater evaluations. The Kolmogorov-Smirnov test was performed at the 95% confidence level to determine whether normality existed. The Wilcoxon (non-parametric test) for related samples was then applied. The Wilcoxon Test (also known as the Mann-Whitney Test) is a test based on rank sums, and is a nonparametric alternative to the two-sample *t* test (Miller and Freund, 1985). This test examines differences in the mean or median of paired observations, with the null hypothesis being that the mean paired differences is 0. If the p-value is small, the idea that difference is due to chance can be rejected and it is safe to conclude that the populations have different medians. If the p-value is large, the overall medians do not differ. P-values less than than 0.05 show significant differences in the medians, while p-values greater than 0.05 reflect that the medians are more similar.

Results illustrated similar behavior in the evaluations of both instructors (see Table 5.9), other than Dimension 6, where the medians were significantly different. For clarity, values less than 0.05 (greater differences in medians) are shaded in blue, while values greater than 0.05 (more similar medians) remain unshaded. Thus, it can be concluded that the designed assembly rubric is homogeneous for accumulative evaluation of CAD assemblies. All statistical results can be found in the appendix.

Table 5.9. Mid-term exam P-Values for Wilcoxon Signed Rank Test.

Oberservation Relationship	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	0.000	0.002	0.276	0.090	0.741	0.000
Individual-Instructor 2	0.000	0.000	0.112	0.446	0.184	0.000
Instructor 1-Instructor 2	0.789	0.059	0.041	0.012	0.011	0.000

5.3 Experiment 2

Following a similar procedure as the first experiement, the final exam required assembling a mechanism. Fifty-one students sat for the exam and submitted self-assessment e-rubrics using the Annota platform. The students were once again assessed on assembly sequence and the use of sub-assemblies. This time, the students were alerted that Dimension 1 (validity) would be assessed as a

"soft" go/no go criterion. As an example, a validity score of 0.5 would result in the other criteria receiving half value.

Students were required to assemble a mechanical filter, using four non-standard parts (previously modeled) and assorted standard parts. The students were again specifically warned on assembly sequence and also on the use of sub-assemblies. Standard parts included an O-ring (DIN 3771, 16 mm ID, 1.8 mm thick), a hex head cap screw (DIN EN 24014, M4 thread, 25 mm long), six round head Allen drive bolts (DIN 7984, M4 thread, 20 mm long, M4 thread, 17.9 mm long), and six M4 hex nuts (ISO 4035 thin). Non-standard parts were provided, except for a purge valve, which the students were required to model.

The preferred assembly strategy was to group the parts based on their function, then group the sub-assemblies. The global group was assembled next, using affordances (assessed in Criterion 6.3) to mate the parts. Non-standard parts included a Cover, Vessel, Case, Spring, Valve Plug, Nozzle, Spinner, Deflector, and Fixation Disk. The assembly can be seen in Figure 5.5.

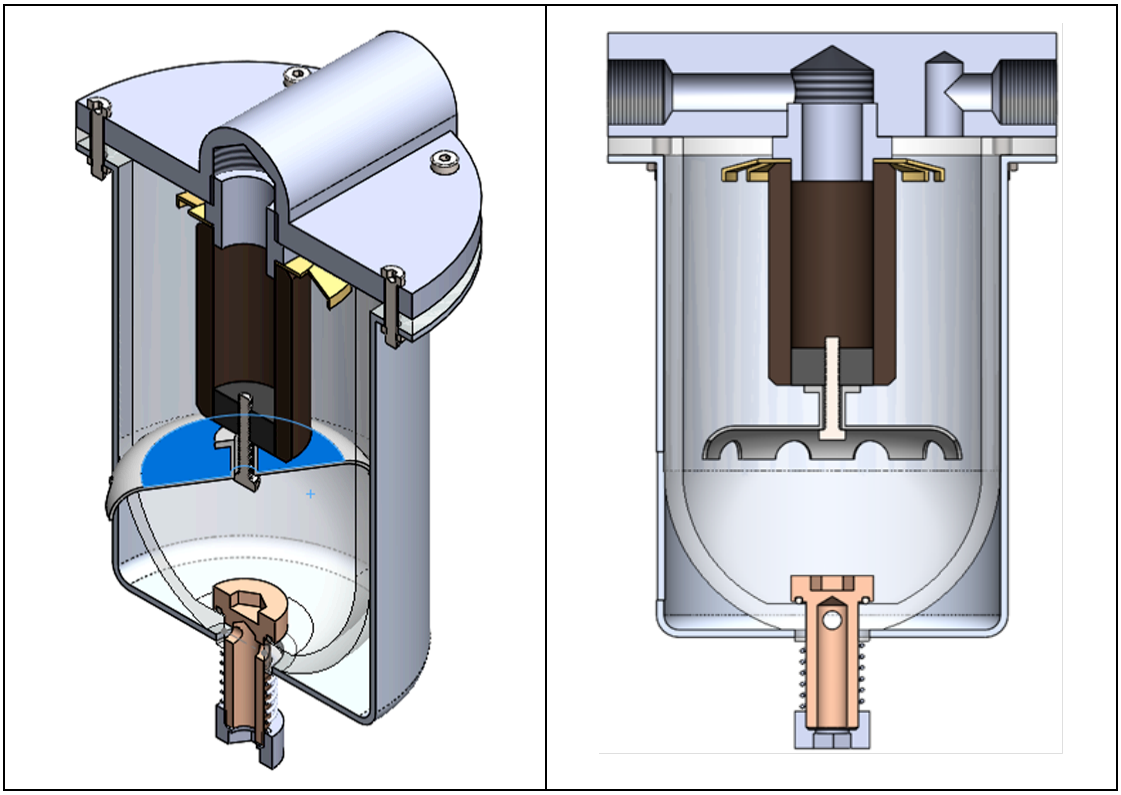


Figure 5.5. Sectioned filter assembly required in Experiment 2. Isometric (left) and Front View (right).

5.4 Experiment 2 Discussion

The main hypothesis to validate in this experiment was that using explicit "soft" go/no go criteria does not affect the correlations, neither between instructor evaluations nor between instructors and students evaluations. The research team searched for differences in the understanding of quality criteria by comparing significant differences between inter-rater evaluations.

Tables 5.10 through 5.15 illustrate the assessment for this task (of each criteria), performed by each student and instructor. As previously shown for the results in Experiment 1, also shown are the instructor average, the difference in scores between instructors, and the difference in scores between the instructor average and each student.

For the dimension of Validity (Table 5.10), both instructors provided exact average scores and the students also gave similar assessments (within 2%). Similar assessments were given for the instructors and students for the dimensions of Completeness (within 7%, Table 5.11), Consistency (within 7%, shown in Table 5.12), and Conciseness (within 5%, shown in Table 5.13). For the dimension of Clarity (Table 5.14), instructors were more optimistic than the students, and the instructors assessments were close. For the dimension of Design Intent (Table 5.15), the instructors were more pessimistic than the students, with Instructor 1 being more pessimistic than Instructor 2.

Table 5.10. Validity scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
2	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
3	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
4	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
6	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
7	0.97	1.00	1.00	1.00	0.03	0.03	0.00	0.03
8	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
9	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
10	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
11	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
12	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
13	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
15	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.89	1.00	1.00	1.00	0.11	0.11	0.00	0.11
18	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
19	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
20	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
23	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
24	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
25	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
26	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
29	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
30	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
31	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
32	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
33	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
34	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
35	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
36	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
37	0.96	1.00	1.00	1.00	0.04	0.04	0.00	0.04
38	0.97	1.00	1.00	1.00	0.03	0.03	0.00	0.03
39	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
40	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
41	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
42	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
43	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
44	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
47	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
48	1.00	0.50	0.50	0.50	-0.50	-0.50	0.00	-0.50
49								
50	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
51	0.90	1.00	1.00	1.00	0.10	0.10	0.00	0.10
52	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
				AVE	-0.02	-0.02	0.00	-0.02

Table 5.11. Completeness scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1-Individual	Δ Instructor 2-Individual	Δ Instructor 1-Instructor 2	Δ Instructor Ave-Individual
1	1.00	0.83	0.83	0.83	-0.17	-0.17	0.00	-0.17
2	0.75	0.75	0.83	0.79	0.00	0.08	-0.08	0.04
3	0.88	0.83	0.92	0.88	-0.04	0.04	-0.08	0.00
4	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
6	0.50	0.83	0.92	0.88	0.33	0.42	-0.08	0.38
7	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
8	1.00	0.83	0.92	0.88	-0.17	-0.08	-0.08	-0.13
9	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
10	0.96	1.00	1.00	1.00	0.04	0.04	0.00	0.04
11	0.50	0.83	0.92	0.88	0.33	0.42	-0.08	0.38
12	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
13	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
14	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
15	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.83	0.83	0.92	0.88	0.00	0.08	-0.08	0.04
18	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
19	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
20	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
23	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
24	1.00	0.58	0.67	0.63	-0.42	-0.33	-0.08	-0.38
25	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
26	0.75	0.42	0.50	0.46	-0.33	-0.25	-0.08	-0.29
27	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
28	0.88	0.83	0.92	0.88	-0.04	0.04	-0.08	0.00
29	0.75	0.83	0.83	0.83	0.08	0.08	0.00	0.08
30	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
31	0.96	0.75	0.75	0.75	-0.21	-0.21	0.00	-0.21
32	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
33	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
34	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
35	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
36	0.83	0.83	0.92	0.88	0.00	0.08	-0.08	0.04
37	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
38	0.88	0.92	1.00	0.96	0.04	0.13	-0.08	0.08
39	0.50	0.92	0.92	0.92	0.42	0.42	0.00	0.42
40	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
41	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
42	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
43	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
44	0.75	0.83	0.83	0.83	0.08	0.08	0.00	0.08
45	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
46	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
47	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
48	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
49								
50	1.00	0.75	0.75	0.75	-0.25	-0.25	0.00	-0.25
51	0.83	0.92	1.00	0.96	0.08	0.17	-0.08	0.13
52	1.00	0.92	1.00	0.96	-0.08	0.00	-0.08	-0.04
				AVE	0.02	0.05	-0.02	0.04

Table 5.12. Consistency scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1-Individual	Δ Instructor 2-Individual	Δ Instructor 1-Instructor 2	Δ Instructor Ave-Individual
1	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
2	0.75	0.00	0.00	0.00	-0.75	-0.75	0.00	-0.75
3	0.69	0.88	0.88	0.88	0.19	0.19	0.00	0.19
4	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
5	1.00	0.38	0.38	0.38	-0.63	-0.63	0.00	-0.63
6	0.50	0.88	1.00	0.94	0.38	0.50	-0.13	0.44
7	0.50	0.50	0.63	0.56	0.00	0.13	-0.13	0.06
8	0.88	0.63	0.63	0.63	-0.25	-0.25	0.00	-0.25
9	0.75	0.88	1.00	0.94	0.13	0.25	-0.13	0.19
10	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
11	0.50	0.38	0.38	0.38	-0.13	-0.13	0.00	-0.13
12	1.00	0.63	0.75	0.69	-0.38	-0.25	-0.13	-0.31
13	1.00	0.75	0.88	0.81	-0.25	-0.13	-0.13	-0.19
14	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
15	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
16	0.75	0.50	0.63	0.56	-0.25	-0.13	-0.13	-0.19
17	0.81	0.88	0.88	0.88	0.06	0.06	0.00	0.06
18	0.81	0.75	0.88	0.81	-0.06	0.06	-0.13	0.00
19	0.88	0.75	0.88	0.81	-0.13	0.00	-0.13	-0.06
20	0.81	0.88	0.88	0.88	0.06	0.06	0.00	0.06
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.50	0.75	0.88	0.81	0.25	0.38	-0.13	0.31
23	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
24	0.75	0.13	0.13	0.13	-0.63	-0.63	0.00	-0.63
25	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
26	0.50	0.38	0.38	0.38	-0.13	-0.13	0.00	-0.13
27	0.75	0.88	0.88	0.88	0.13	0.13	0.00	0.13
28	0.69	0.75	0.88	0.81	0.06	0.19	-0.13	0.13
29	1.00	0.25	0.38	0.31	-0.75	-0.63	-0.13	-0.69
30	0.63	0.63	0.75	0.69	0.00	0.13	-0.13	0.06
31	1.00	0.00	0.13	0.06	-1.00	-0.88	-0.13	-0.94
32	0.75	0.75	0.88	0.81	0.00	0.13	-0.13	0.06
33	0.75	0.50	0.63	0.56	-0.25	-0.13	-0.13	-0.19
34	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
35	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
36	0.81	0.88	0.88	0.88	0.06	0.06	0.00	0.06
37	0.88	0.63	0.75	0.69	-0.25	-0.13	-0.13	-0.19
38	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
39	0.75	0.50	0.63	0.56	-0.25	-0.13	-0.13	-0.19
40	0.75	0.88	0.88	0.88	0.13	0.13	0.00	0.13
41	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
42	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
43	0.88	0.63	0.75	0.69	-0.25	-0.13	-0.13	-0.19
44	0.88	0.50	0.63	0.56	-0.38	-0.25	-0.13	-0.31
45	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
46	0.75	0.50	0.63	0.56	-0.25	-0.13	-0.13	-0.19
47	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
48	1.00	0.63	0.75	0.69	-0.38	-0.25	-0.13	-0.31
49								
50	0.75	0.00	0.13	0.06	-0.75	-0.63	-0.13	-0.69
51	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
52	0.88	0.63	0.75	0.69	-0.25	-0.13	-0.13	-0.19
				AVE	-0.11	-0.04	-0.07	-0.07

Table 5.13. Conciseness scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
2	0.25	0.58	0.58	0.58	0.33	0.33	0.00	0.33
3	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
4	0.50	0.67	0.67	0.67	0.17	0.17	0.00	0.17
5	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
6	0.50	0.83	0.83	0.83	0.33	0.33	0.00	0.33
7	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
8	1.00	0.58	0.75	0.67	-0.42	-0.25	-0.17	-0.33
9	0.75	0.33	0.67	0.50	-0.42	-0.08	-0.33	-0.25
10	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
11	0.50	0.92	0.92	0.92	0.42	0.42	0.00	0.42
12	1.00	0.58	0.58	0.58	-0.42	-0.42	0.00	-0.42
13	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
14	0.50	0.33	0.33	0.33	-0.17	-0.17	0.00	-0.17
15	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
16	1.00	0.58	0.67	0.63	-0.42	-0.33	-0.08	-0.38
17	0.46	1.00	1.00	1.00	0.54	0.54	0.00	0.54
18	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
19	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
20	0.67	0.83	0.83	0.83	0.17	0.17	0.00	0.17
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.50	0.58	0.58	0.58	0.08	0.08	0.00	0.08
23	0.50	0.67	0.67	0.67	0.17	0.17	0.00	0.17
24	0.75	0.33	0.42	0.38	-0.42	-0.33	-0.08	-0.38
25	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
26	0.75	0.33	0.33	0.33	-0.42	-0.42	0.00	-0.42
27	0.50	0.92	0.92	0.92	0.42	0.42	0.00	0.42
28	0.67	1.00	1.00	1.00	0.33	0.33	0.00	0.33
29	0.75	0.58	0.58	0.58	-0.17	-0.17	0.00	-0.17
30	0.67	0.83	0.83	0.83	0.17	0.17	0.00	0.17
31	0.92	0.67	0.67	0.67	-0.25	-0.25	0.00	-0.25
32	0.75	0.83	0.83	0.83	0.08	0.08	0.00	0.08
33	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00
34	0.75	0.83	0.83	0.83	0.08	0.08	0.00	0.08
35	1.00	0.92	0.92	0.92	-0.08	-0.08	0.00	-0.08
36	0.67	0.42	0.50	0.46	-0.25	-0.17	-0.08	-0.21
37	0.79	1.00	1.00	1.00	0.21	0.21	0.00	0.21
38	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
39	0.50	0.42	0.42	0.42	-0.08	-0.08	0.00	-0.08
40	1.00	0.58	0.92	0.75	-0.42	-0.08	-0.33	-0.25
41	0.83	0.92	0.92	0.92	0.08	0.08	0.00	0.08
42	0.92	1.00	1.00	1.00	0.08	0.08	0.00	0.08
43	0.83	0.83	0.83	0.83	0.00	0.00	0.00	0.00
44	0.67	0.83	0.83	0.83	0.17	0.17	0.00	0.17
45	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
46	0.50	0.83	0.83	0.83	0.33	0.33	0.00	0.33
47	0.50	0.92	0.92	0.92	0.42	0.42	0.00	0.42
48	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
49								
50	0.50	0.75	0.75	0.75	0.25	0.25	0.00	0.25
51	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
52	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
				AVE	0.04	0.07	-0.02	0.05

Table 5.14. Clarity scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1- Individual	Δ Instructor 2- Individual	Δ Instructor 1- Instructor 2	Δ Instructor Ave- Individual
1	0.75	0.63	0.75	0.69	-0.13	0.00	-0.13	-0.06
2	0.25	0.50	0.50	0.50	0.25	0.25	0.00	0.25
3	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
4	0.94	1.00	1.00	1.00	0.06	0.06	0.00	0.06
5	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
6	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
7	0.88	0.75	0.88	0.81	-0.13	0.00	-0.13	-0.06
8	0.63	0.25	0.50	0.38	-0.38	-0.13	-0.25	-0.25
9	0.50	0.88	0.88	0.88	0.38	0.38	0.00	0.38
10	0.88	0.88	1.00	0.94	0.00	0.13	-0.13	0.06
11	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
12	1.00	0.50	0.63	0.56	-0.50	-0.38	-0.13	-0.44
13	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
14	0.50	0.25	0.38	0.31	-0.25	-0.13	-0.13	-0.19
15	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
16	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
17	0.31	0.38	0.38	0.38	0.06	0.06	0.00	0.06
18	0.85	0.75	0.75	0.75	-0.10	-0.10	0.00	-0.10
19	0.52	0.75	0.75	0.75	0.23	0.23	0.00	0.23
20	0.63	0.63	0.75	0.69	0.00	0.13	-0.13	0.06
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.50	0.88	0.88	0.88	0.38	0.38	0.00	0.38
23	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
24	1.00	0.88	0.88	0.88	-0.13	-0.13	0.00	-0.13
25	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
26	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
27	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
28	0.83	1.00	1.00	1.00	0.17	0.17	0.00	0.17
29	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
30	0.75	0.88	0.88	0.88	0.13	0.13	0.00	0.13
31	0.50	0.75	0.75	0.75	0.25	0.25	0.00	0.25
32	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
33	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
34	1.00	0.88	0.88	0.88	-0.13	-0.13	0.00	-0.13
35	1.00	0.88	0.88	0.88	-0.13	-0.13	0.00	-0.13
36	0.65	0.63	0.63	0.63	-0.02	-0.02	0.00	-0.02
37	0.75	0.38	0.50	0.44	-0.38	-0.25	-0.13	-0.31
38	0.71	1.00	1.00	1.00	0.29	0.29	0.00	0.29
39	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
40	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
41	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
42	0.63	0.75	0.75	0.75	0.13	0.13	0.00	0.13
43	0.63	0.63	0.63	0.63	0.00	0.00	0.00	0.00
44	0.75	0.88	0.88	0.88	0.13	0.13	0.00	0.13
45	0.75	1.00	1.00	1.00	0.25	0.25	0.00	0.25
46	0.75	0.63	0.63	0.63	-0.13	-0.13	0.00	-0.13
47	0.50	1.00	1.00	1.00	0.50	0.50	0.00	0.50
48	0.75	0.50	0.50	0.50	-0.25	-0.25	0.00	-0.25
49								
50	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
51	0.81	1.00	1.00	1.00	0.19	0.19	0.00	0.19
52	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
				AVE	0.09	0.12	-0.02	0.10

Table 5.15. Design Intent scores for final exam.

Student	Individual Score	Instructor 1 Score	Instructor 2 Score	Instructor Average	Δ Instructor 1-Individual	Δ Instructor 2-Individual	Δ Instructor 1-Instructor 2	Δ Instructor Ave-Individual
1	0.50	0.58	0.75	0.67	0.08	0.25	-0.17	0.17
2	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
3	0.71	0.67	0.75	0.71	-0.04	0.04	-0.08	0.00
4	0.67	0.75	0.83	0.79	0.08	0.17	-0.08	0.12
5	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
6	0.50	0.58	0.75	0.67	0.08	0.25	-0.17	0.17
7	0.71	0.67	0.83	0.75	-0.04	0.13	-0.17	0.04
8	0.75	0.25	0.33	0.29	-0.50	-0.42	-0.08	-0.46
9	0.50	0.50	0.67	0.58	0.00	0.17	-0.17	0.08
10	0.88	1.00	1.00	1.00	0.13	0.13	0.00	0.13
11	0.75	0.50	0.58	0.54	-0.25	-0.17	-0.08	-0.21
12	1.00	0.75	0.75	0.75	-0.25	-0.25	0.00	-0.25
13	0.75	0.83	0.92	0.88	0.08	0.17	-0.08	0.13
14	0.75	0.25	0.42	0.33	-0.50	-0.33	-0.17	-0.42
15	0.75	0.58	0.67	0.63	-0.17	-0.08	-0.08	-0.13
16	0.92	0.92	0.92	0.92	0.00	0.00	0.00	0.00
17	0.33	0.33	0.42	0.38	0.00	0.08	-0.08	0.04
18	0.79	0.42	0.58	0.50	-0.38	-0.21	-0.17	-0.29
19	0.79	0.42	0.50	0.46	-0.38	-0.29	-0.08	-0.33
20	0.46	0.17	0.42	0.29	-0.29	-0.04	-0.25	-0.17
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.50	0.75	0.75	0.75	0.25	0.25	0.00	0.25
23	0.50	0.42	0.50	0.46	-0.08	0.00	-0.08	-0.04
24	1.00	0.08	0.17	0.13	-0.92	-0.83	-0.08	-0.88
25	0.67	0.75	0.75	0.75	0.08	0.08	0.00	0.08
26	0.25	0.33	0.33	0.33	0.08	0.08	0.00	0.08
27	0.50	0.50	0.67	0.58	0.00	0.17	-0.17	0.08
28	0.79	0.50	0.75	0.63	-0.29	-0.04	-0.25	-0.17
29	1.00	0.42	0.50	0.46	-0.58	-0.50	-0.08	-0.54
30	0.50	0.50	0.58	0.54	0.00	0.08	-0.08	0.04
31	0.88	0.25	0.42	0.33	-0.63	-0.46	-0.17	-0.54
32	0.50	0.42	0.50	0.46	-0.08	0.00	-0.08	-0.04
33	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00
34	1.00	0.58	0.67	0.63	-0.42	-0.33	-0.08	-0.38
35	0.75	0.50	0.75	0.63	-0.25	0.00	-0.25	-0.13
36	0.54	0.42	0.50	0.46	-0.13	-0.04	-0.08	-0.08
37	0.79	0.75	0.75	0.75	-0.04	-0.04	0.00	-0.04
38	0.79	0.75	0.83	0.79	-0.04	0.04	-0.08	0.00
39	0.75	0.67	0.67	0.67	-0.08	-0.08	0.00	-0.08
40	0.75	0.42	0.50	0.46	-0.33	-0.25	-0.08	-0.29
41	0.83	0.50	0.67	0.58	-0.33	-0.17	-0.17	-0.25
42	0.83	0.67	0.75	0.71	-0.17	-0.08	-0.08	-0.12
43	0.58	0.25	0.42	0.33	-0.33	-0.17	-0.17	-0.25
44	0.58	0.42	0.58	0.50	-0.17	0.00	-0.17	-0.08
45	0.75	0.92	0.92	0.92	0.17	0.17	0.00	0.17
46	0.50	0.50	0.58	0.54	0.00	0.08	-0.08	0.04
47	0.50	0.83	0.83	0.83	0.33	0.33	0.00	0.33
48	0.75	0.42	0.50	0.46	-0.33	-0.25	-0.08	-0.29
49								
50	0.50	0.42	0.50	0.46	-0.08	0.00	-0.08	-0.04
51	0.67	0.75	0.83	0.79	0.08	0.17	-0.08	0.12
52	0.88	0.67	0.75	0.71	-0.21	-0.13	-0.08	-0.17
				AVE	-0.14	-0.05	-0.09	-0.10

Table 5.16 illustrates the inter-rater reliability scores for the final exam (for the student and both instructors). It is shown that once again, there is greater agreement between the instructors than between instructor and students. There exists moderate to strong agreement for Dimension 1, between both instructors and between instructors and students. There is strong agreement between instructors for Dimensions 1, 2, 4, and 5 and little agreement between instructors and students for any dimension other than validity. It appears that there is no measureable increase in agreement for all dimensions other than validity, for instructors and students, between the mid-term and final exam. Reasons for the lack of increase could be that there was only three weeks between exams, giving little time for the students to grasp missed concepts in order to improve their performance.

Table 5.16. Inter-rater reliabilty scores for final exam

% of Agreement	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	69.0	34.6	23.0	15.0	23.0	15.0
Individual-Instructor 2	69.0	30.0	15.0	15.0	23.0	15.0
Instructor 1-Instructor 2	100.0	75.0	44.0	88.0	84.6	25.0

Table 5.17 shows the Pearson Correlation values for the final exam (for the students and both instructors). It can be seen that there is high correlation between instructors for all dimensions (increased correlation for all dimensions, except a slight decrease in Dimension 4, but still exhibiting strong correlation). As with the mid-term exam, there is less correlation between instructors and students. Explicitly, there is strong to moderate correlation between instructors and students for Dimensions 1, 2, and 5, but low correlation for Dimensions 3 and 4. Dimension 4 appears unchanged.

Table 5.17. Pearson Correlation values for final exam.

Correlation Coefficient	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	0.73822	0.73258	0.31092	0.48536	0.54167	0.43169
Individual-Instructor 2	0.73822	0.73512	0.35272	0.5751	0.60329	0.48355
Instructor 1-Instructor 2	1	0.98461	0.97757	0.96389	0.98004	0.95149

Table 5.18 shows the p-values for the Wilcoxon Signed Rank Test, as explained earlier in the chapter. Once again, for clarity, values less than 0.05 (greater differences in medians) are shaded in blue, values greater than 0.05 (more similar medians) remain unshaded, and values that could not be computed are shaded in grey. In this case, for Dimension 1, the medians were not significantly different, enough that the p-value could not be computed between Instructor 1 and Instructor 2.

Table 5.18. Final exam P-Values for Wilcoxon Signed Rank Test.

Oberservation Relationship	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	1.000	0.231	0.023	0.239	0.007	0.000
Individual-Instructor 2	1.000	0.047	0.722	0.047	0.000	0.194
Instructor 1-Instructor 2	Cannot Compute	0.002	0.000	0.036	0.014	0.000

Table 5.19 shows whether each dimension increased, decreased, or remained unchanged in respect to rubric understanding, between the mid-term and final exams for inter-rater reliability. As can be seen, Dimensions 1, 5, and 6 displayed increased rater agreement, but a decrease in understanding is shown for Dimensions 2, 3, and 4.

Table 5.19. Percent Agreement values between mid-term and final exam.

Difference	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	Increase	Decrease	Decrease	Decrease	Same	Increase
Individual-Instructor 2	Increase	Decrease	Decrease	Decrease	Increase	Increase
Instructor 1-Instructor 2	Increase	Same	Decrease	Increase	Increase	Decrease

Table 5.20 reflects whether each dimension increased, decreased, or remained unchanged in respect to rubric understanding, between the mid-term and final exams for the Pearson Correlation values. Dimensions 1, 2, 4, and 5 showed increased correlation, but decreased correlation is reported for Dimensions 3 and 6.

Table 5.20. Pearson Correlation values between mid-term and final exam.

Difference	Dimension 1 (Valid)	Dimension 2 (Complete)	Dimension 3 (Consistent)	Dimension 4 (Concise)	Dimension 5 (Clear)	Dimension 6 (Design Intent)
Individual-Instructor 1	Increase	Increase	Decrease	Increase	Increase	Decrease
Individual-Instructor 2	Increase	Increase	Decrease	Increase	Increase	Decrease
Instructor 1-Instructor 2	Increase	Increase	Increase	Decrease	Increase	Increase

Ideally, it would be useful to determine if the correlation for each dimension improved or decreased in a significant manner, but since the r-value is synthetically bound between 0 and 1, it is exceedingly difficult to construct meaningful conclusions about this matter. A linear relationship cannot be assumed between the correlation values, but even if the change in correlation values were significant, would it be consequential? Even with perfectly defined rubric dimensions, it is impossible to remove all subjectivity, which clouds any definitive judgment. In such cases, only the professional expertise of the investigator would guide those determinations. Regardless of this lack of statistical certainty, a pronounced general pattern emerges that reflects a positive directional improvement for a majority of rubric dimensions (between both instructor and student, and between instructors).

In order to calibrate the assembly rubric more fully (and to gain desired statistical significance), additional steps should be taken in the future. One such improvement could be conducting an experiment where students are provided with identical assembly models (with separate trials examining models constructed at varying quality levels) and have them assess these models. The models could then be compared against an ideal solution provided by an instructor (or group of instructors). This experiment would provide ample degrees of freedom (by furnishing multiple observations of the same event) in which to perform various statistical tests (ex. Paired t-test) and would theoretically remove any assessment bias that students may exhibit toward their own models.

6. Conclusions and Recommendations for Future Work

It can be seen from these two experiments that there is greater agreement and correlation between the instructors than between the instructors and students, for all rubric dimensions. There is strong to moderate correlation between the instructors for the dimensions of validity, completeness,

conciseness, and clarity. Little correlation exists for the dimensions of consistency and design intent. Probable reasons for the lack of correlation for these two dimensions could be that they are more complex modeling concepts, consisting of intimate knowledge of:

- Position of the model in reference to various reference systems.
- Understanding of proper and improper movement of components within the assembly.
- Purpose of the mechanism, how it functions, and which components are needed as anchors within the overall model.

In summary, the small differences between instructors allows the conclusion that the proposed assemblies rubric is mature enough to provide an objective accumulative assessment of students. Thus, it can be determined that raters can be safely used interchangeably.

However, it still has limited validity to provide formative self-evaluation of CAD assembly skills for CAD trainees. The research team surmises that improving the instructional materials (primarily in the dimensions of consistency and design intent) is perhaps, the first future step to obtain valid formative self-evaluation of CAD assembly skills for beginning users.

Finally, it appears that the required improvements do not primarily depend on small improvements (such as introducing soft go/no go criterion), or on a moderate increase in the exposure to rubrics.

Chapter 6
Conclusions and Future Work

1. Summary of Findings

This dissertation proposed four objectives (presented in Chapter 1) and they are restated below, with the result and fulfilment of each.

1.1 Objective 1: Literature Review

The first goal was accomplished by examining previous research in regards to the concept of CAD model quality, in order to determine how the dimensions of quality models can be detailed in suitable assertions, how good modeling practices should be organized in a new set of rubrics, while making CAD model quality the main goal.

An exhaustive survey was conducted on how CAD model quality has been historically defined, most often described by examining design intent, with its historical relationship with design rationale. Although design rationale has been well-established in industrial settings, in the context of MCAD systems, design intent remains a complex and misunderstood concept, which is reflected in the results of various experiments conducted and explained in this thesis.

According to the literature review (extensively described in Chapter 2), it can be concluded that design intent is frequently described as a model's anticipated behavior once it undergoes alteration, and there is acknowledgement that modeling approaches influence design intent communication. There are significant advantages to express design intent through suitable modeling strategies, especially for CAD beginners.

Strategies designed to improve the communication of design intent in CAD models to enhance their quality, jointly with guidelines targeted to evaluate efficiency, are currently receiving notice and have been addressed with this research. It is apparent that metrics directed toward the instruction of design intent are necessary, since it has been previously stated that design intent transmission transferred through CAD models may be performed at three different stages (sketches, reference datums, and the modeling operations themselves), which have competing tradeoffs which must be balanced to arrive at the best modeling strategy.

Previous research also reflects that rubrics are a valuable device to expedite consistent design intent communication. Rubrics are vital not only for evaluation, but also for the communication of instructor expectations. This research examined how to more clearly define qualities of design intent (among other rubric dimensions) to enable easier CAD model assessment.

The new contributions in regards of the literature review were compiled and published as two conference papers and journal article:

- Otey, J., Company, P., Contero, M., & Camba, J. D. (2014, June). A review of the design intent concept in the context of CAD model quality metrics. *ASEE Annual Conference and Exposition,* Indianapolis, IN. http://www.asee.org/conferences-and-events/conferences/annual-conference/2014. Paper ID #9556.

The role of the canditate in this publication was primary authorship, literature review, creation of presentation, and delivery.

- Camba J., Contero M., Otey J., Company P. (2014, August). Explicit Communication of Geometric Design Intent in CAD: Evaluating Annotated Models in the Context of Reusability. Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference DETC/DTM 2014. August 17-20, 2014, Buffalo, New York, US. 26th International Conference on Design Theory and Methodology (DTM) Paper DETC2014-34527.

This paper acceptance criteria at this conference was competitive, with an acceptance rate of fifty-eight percent. The role of the canditate in this publication was editing, creation of presentation material, and delivery.

- Otey J., Company P., Contero M. and Camba J.D. (2017). Revisiting the Design Intent Concept in the Context of Mechanical CAD Education. *Computer-Aided Design and Applications* (ISSN 1686-4360). 15(1), 2018 (DOI: 10.1080/16864360.2017.1353733).

The role of the canditate in this publication was primary authorship, literature review, and corresponding author.

1.2 Objective 2: Rubric Development

The second objective (detailed in Chapters 3 and 5) can be briefly described as developing an original set of rubrics, created and designed (with associated graphical assertions maps), to emphasize proper modeling practices while stressing the importance of creating robust models through proper design intent, so that the corresponding assemblies can be updated with minimal effort.

Care was exercised to preserve consistency between the parts and assembly rubrics, with variation existing between the descriptors and achievement levels for each criterion, which were restated to be more applicable to assemblies. The use of ambiguous and explicit descriptors were required to prevent homogeneous evaluation and for proper evaluation to occur. Achievement levels reflected the amount of conformity for each assessed aspect, with simple cases being dichotomously determined.

The rubric generated scores, as scoring is consubstantial to the rubric. The scoring process was defined with formulas, to provide the aggregated score from the achievement levels. The rubric was defined by three characteristics: dichotomous criterion, evaluation criteria with varying importance levels, and Go/No Go criteria. Rubric dimensions were divided into six categories: valid, complete, consistent, concise, clear, and design intent. Annota e-Rubrics were used during the experiments.

The new contributions in regards of the validation of rubrics have been thus far, compiled and published as one conference paper and two journal articles:

- Company P., Otey J., Camba J., Contero M. (2014, August). Leveraging Mechanical 3D CAD Systems Through Improved Model Quality Based on Best Practices and Rubrics. Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference DETC/DTM 2014 August 17-20, 2014, Buffalo, New York, US. 26th International Conference on Design Theory and Methodology (DTM) Paper no. DETC2014-34489.

This paper acceptance criteria at this conference was competitive, with an acceptance rate of fifty-eight percent. The role of the canditate in this publication was secondary authorship, editing, creation of presentation material, and delivery.

- Company P., Contero M., Otey J. and Plumed R. (2015). Approach for developing coordinated rubrics to convey quality criteria in MCAD training. *Computer-Aided Desi*gn (ISSN 0010-4485). Vol. 63, (2015), 101–117.

The canditate was a member of the research team that developed a set of rubrics to promote quality of CAD models. Subsequently, the candidate was the leader in the development of a similar set of rubrics to promote quality of CAD assemblies, although this new contribution is now being prepared for journal submission.

- Company P., Otey J., Contero M., Agost M.J. and Almiñana A. (2016). A. Implementation of Adaptable Rubrics for CAD Model Quality Formative Assessment. *International Journal of Engineering Education* (ISSN 0949-149X). Vol 32(2A). pp. 749–761. 2016.

The role of the canditate in this publication was secondary authorship and editing.

1.3 Objective 3: Rubric Validation

The task developed to accomplish the third objective (detailed in Chapters 4 and 5), can be summarized as test and statistically validate these rubrics in a classroom setting, so that objective determinations can be made about the link between specific procedures and the quality of the corresponding created model.

It has been shown from the conducted experiments that there is more inter-rater agreement and correlation between instructors than between instructors and students, for all rubric dimensions. There is strong to moderate correlation between instructors for the dimensions of validity, completeness, conciseness, and clarity, while slight correlation exists for the dimensions of consistency and design intent. Likely reasons for the lack of correlation for these two dimensions might be that they are complex modeling concepts entailing of familiar comprehension of:

- Model position in reference to various reference systems.
- Understanding of both proper and improper component movement within the assembly.
- Purpose of the mechanism, how it functions, and the components needed as anchors within the overall model.

The minor differences between instructors allows the conclusion that the proposed assemblies rubric provides an objective accumulative assessment of students. Thus, it can be assumed that raters can be safely interchanged.

The new contributions in regards of the validation of rubrics have been thus far, compiled and published as a conference paper:

- Company P., Otey J., Camba J., Contero M. (2014, August). Leveraging Mechanical 3D CAD Systems Through Improved Model Quality Based on Best Practices and Rubrics. Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference DETC/DTM 2014 August 17-20, 2014, Buffalo, New York, US. 26th International Conference on Design Theory and Methodology (DTM) Paper no. DETC2014-34489.

This paper acceptance criteria at this conference was competitive, with an acceptance rate of fifty-eight percent. The canditate was a member of the research team that developed a set of rubrics to promote quality of CAD models. Subsequently, the candidate was the leader in the development of a similar set of rubrics to promote quality of CAD assemblies, although this new contribution is now being prepared for journal submission.

1.4 Objective 4: Development of Pedagogical Tools

The task developed to accomplish the third objective (detailed in Chapter 4), can be summarized as introducing tools, from early in the instruction, so proper modeling practices can be instilled (while also eliminating poor practices), so poorly created part models do not negatively affect corresponding assembly models.

Assertions map were developed, illustrating how the expand-contract strategy adapts the rubrics to CAD trainee progress and assists in understanding the different rubric dimensions. The expand-contract process is advantageous to comprehend the quality concepts embedded in the rubrics, and the students require instructor evaluations in order to compare their own work against an ideal solution to fully understand the rubric dimensions. Detailed rubric dimension curricular material (consisting of checklists, good practices, and various evaluation tools) are advantageous for student understanding, especially when considering trade-offs to achieve valid designs. Rubrics need to be adaptable (modified by the user), which should be easily understood and user-friendly, and adaptive (rubric can change itself, depending on the usage pattern).

The new contributions in regards of the development of pedagogical tools have been so far compiled and published as two journal papers:

- Company P., Otey J., Contero M., Agost M.J. and Almiñana A. (2016). Implementation of Adaptable Rubrics for CAD Model Quality Formative Assessment. *International Journal of Engineering Education* (ISSN 0949-149X). Vol 32(2A). pp. 749–761. 2016.

The role of the canditate in this publication was secondary authorship and editing.

- Company P. Contero M., Otey J., Camba J.D., Agost M.J. and Pérez-López D.C. (2017). Web-based system for adaptable rubrics: case study on CAD assessment. *Journal of Educational Technology & Society* (ISSN 1436-4522 (online) and 1176-3647 (print)). 20(3), pp.24-41. 2017.

The role of the canditate in this publication was secondary authorship, experimental assessment, and data analysis.

2. Conclusions

2.1 Rubrics

Educational rubrics are curricular devices used to assess student performance, oftentimes displayed as a matrix comprised of various measureable criteria, which are then evaluated through achievement

levels. Rubrics can exist in both evaluative and formative formats. Evaluative rubrics are used when an expert determines the pedagogical progress of a learner, while formative rubrics are employed by the learners themselves, in order to chart their progress and identify scholastic deficiencies that are in need of remediation.

Rubrics must be continually refined and improved, in an iterative, collaborative process, until satisfactory agreement is attained, both between raters, but also between raters and learners. Students must be convinced of the benefits of using rubrics, or their reluctance to use them will skew their scores and make them futile. Oftentimes, students are tempted to take shortcuts in engineering graphics, as when they skip the "sketching" step of the design process in order to advance directly to CAD. For those students, it can be anticipated that the use of rubrics will be perceived to be a tedious, unnecessary step, preventing them from quickly moving toward task completion. They need to be persuaded of the utility of using rubrics.

Rubrics can also be described as being either static or dynamic. Static rubrics, existing in paper form only, do not provide immediate feedback to the learner. Dynamic rubrics, perform calculations that provide immediate evaluative observations to the user, but also can be independently adapted to situations, depending on the capability of the user. Electronic rubrics (eRubrics) are ideally suited for dynamic rubrics, and permit the use and development of both adaptable and adaptive rubrics.

In conclusion, based on the assembly rubric experiments described in Chapter 5, it is apparent that the modest differences between instructors suggests that the proposed assemblies rubric is sufficiently sophisticated to furnish an unbiased accumulative assessment of student performance. Accordingly, it can be confidently stated that raters can be used interchangeably without sacrificing accuracy.

However, the assembly rubric possesses finite efficacy to produce formative self-evaluation of CAD assembly skills for new learners. The improvement of instructional materials (principally for consistency and design intent) is suggested as the starting point of obtaining authoritative formative self-evaluation of CAD assembly skills.

Of course, this ambit is not without difficulty, specifically for the dimension of design intent. Of all of the rubric dimensions, the meaning of design intent has proven to be the most difficult to not only define, but to convey to the students. There are many reasons for this to be so, but primarily, design intent recurrently requires precise prior knowledge of how the mechanism will and should perform, an awareness that may be beyond the comprehension level of inexperienced users. This inexperience is not only grounded in a lack of understanding of how to properly use the software, but oftentimes relates to a student's lack of real-world cognizance. Lastly, it occurs that needed enhancements do not solely rely on minor, trivial developments or on a modest increase in rubric exposure.

3. Future Work

3.1 Rubric Refinement

As stated in the previous section, further refinement is necessary to attain assessment accuracy, primarily for the dimensions of consistency and design intent. More easily understood introductory material for students could possibly increase comprehension of various misunderstood rubric dimensions, but perhaps, as also previously stated, these concepts are beyond beginner level. Perhaps it would be preferable to refine the assembly rubric for more advanced students (from a formative perspective), or to develop an advanced version of the rubric (from an evaluative standpoint) to be more easily used and adopted by experienced raters.

3.2 Rubric Tools

Parts and assembly rubrics, especially if offered in electronic formats, offer the possibility of them being integrated into tools which could be incorporated into the CAD software interface. These tools could be used to not only guide the CAD operator in the design process, but could also assess the model undergoing creation. Plugins allowing real-time collaboration have been created in the past to integrate with modeling software (Camba et al. 2014c), so this ambit definitely merits further examination.

3.3 Automating Rubrics

The use of e-Rubrics presents a multitude of opportunities for creating linkages between the rubric and CAD software, facilitating the automation of the assessment process. An increase in dichotomous criteria would make automation significantly easier, thus some deficiencies (such as validity, consistency or completeness) could be identified and remedied by the modeling software. Automating the rubric process will curtail manual input, increasing the likelihood of student use, especially for those who have expressed hesitation in the past (in using paper-based rubrics). This automation would benefit not only educational settings, but also industry, where files are created and shared by multiple design team members, oftentimes in different geographic locations, and the difference between using a high quality model versus a medium/low quality model results in a significant reduction of time-to-market and money.

3.4 Rubrics in Education

Any adoption of e-Rubrics, or the automation of them, could provide immense benefits to the future of CAD training, especially in regards to their use in on-line or distance education. One of the main impediments to on-line engineering graphics education is the difficulty of, and labor-intensive processes required, to provide timely and efficient student feedback. Any attempts to improve this process would return great dividends to the entire educational process. Additionally, with rubric automation, students could receive instantaneous evaluation prior to task submittal, encouraging self-assessment and improved scores, thereby resulting in student ownership of their training. This process would produce increased confidence among the students and in their design skills.

Any increase in automated rubrics would provide for an increased efficiency in educational content delivery, thereby allowing instructors to serve more students. This increased efficiency could be manifested in more consistent task evaluations, reduced course completion time for those institutions that offer a “learn-at-your-own-pace” scholastic terms, or courses that are only offered in off-times (such as mini-mesters occurring over the holidays).

3.5 Drawings Rubrics

Furthermore, an obvious next step would be to create rubrics to assist in assessment of parametrically created drawings. Of course, abundant literature already exists that define proper methods to create mechanical drawings (ex. ANSI 14.5M, ISO 128, ASME Y14.41-2012, ASME Y14.100-2013, etc.), so use of these standards would be an apparent starting point for any drawing rubric development. Possibly the current dimension titles (valid, concise, etc.) could be re-purposed for the drawing rubric, to remain consistent with previous rubrics developed by the research team. However, it must be noted that creating drawings from scratch is different from *extracting* drawings from parametric models (since extracted files are linked to their correspondimg models).

Problems present themselves in such ways that in many cases, an increase in the use of geometric relationships in the part file reduces the number of “driving” dimensions from the model automatically imported into the drawing file. In those situations, additional “driven” dimensions are required to

properly annotate the model and ensure adequate design communication. And if design intent is often understood to be how a model can be updated, can that information be ascertained from a drawing in the first place? In any event, attempts should be made regardless, as any step is a move toward better understanding.

3.6 Rubrics for Hybrid Surface and Direct Modeling

While this research has focused on feature-based modeling, new developments in Hybrid Surface-Modeling and Direct Modeling open up new avenues for rubric exploration and development. Feature-based modeling provides exact control of the geometry for each feature of an artifact and is ideally suited for manufacturing environments where multiple variations of each part can be quickly obtained. Conversely, using software with such rigorous task processes may inhibit creativity in the early stages of design exploration, where direct manipulation of the part could be beneficial. Rubrics could of course be constructive to these new modeling scenarios, and feature-based rubrics would need to be modified for them to be useful, but the overlying principles maintain applicability.

4. Concluding Remarks

Appraisal of student performance is a critical component necessary for engaged student learning. The use of rubrics to perform assessment not only serves as a method for instructors to objectively judge student work, but also can provide important learner self-assessment in order for the students to develop ownership of their own training. This study examined the use of assembly rubrics, described how they evolved from parts rubrics developed by the same research team, and studied how they affect student self-evaluation of their CAD assembly skills. In addition, instructor assessment of students was evaluated. While the assembly rubrics were partially understood and effectively used by the students, the assembly rubric was more clearly understood and utilized by the instructors.

REFERENCES

Allsop, C. V. P. (2009). *The development of three-dimensional Computer Aided Design (CAD) modeling strategies and an investigation into their impact on novice users* (Doctoral dissertation). Retrieved from Loughborough University Institutional Repository. (https://dspace.lboro.ac.uk/2134/5398).

Amadori, K., Tarkian, M., Ölvander, J., & Krus, P. (2012). Flexible and robust CAD models for design automation. *Advanced Engineering Informatics, 26(2),* 180–95.

Anderl, R., & Mendgen, R. (1998). Analyzing and optimizing constraint-structures in complex parametric CAD models. In: B. Bruderlin & D. Roller (Eds.), *Geometric constraints and applications* (pp.58-81). Berlin: Springer.

Anderson, B. & Ansaldi, S. (1998, September). *ENGEN Data Model: a neutral model to capture design intent*. Presented at the Tenth International IFIP WG Conference, Trento, Italy.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation, 10(3)*, 1-11. http://pareonline.net/getvn.asp?v=10&n=3.

Anglin, L., Anglin, K., Schuman, P. L., & Kaliski, J. A. (2008). Improving the efficiency and effectiveness of grading through the use of computer-assisted grading rubrics. *Decision Sciences Journal of Innovative Education, 6(1)*, 51-73.

Anthony, L., Regli, W. C., John, J. E., & Lombeyda, S. V. (2001). An approach to capturing structure, behavior, and function of artifacts in computer-aided design. *Journal of Computing and Information Science in Engineering, 1(2),* 186-192.

Atkinson, D., & Lim, S. L. (2013). Improving assessment processes in higher education: Student and teacher perceptions of the effectiveness of a rubric embedded in a LMS. *Australasian Journal of Educational Technology, 29(5),* 651-666.

Ault, H. K. (1999). Using geometric constraints to capture design intent. *Journal for Geometry and Graphics, 3(1),* 39-45.

Ault, H. K. (1999). 3-D geometric modeling for the 21st Century. *Engineering Design Graphics Journal, 63(2),* 33-42.

Ault, H. K. & Giolas, D. T. (2005). An investigation of solid modeling practices in industry. *Engineering Design Graphics Journal, 69(1),* 34-43.

Auvinen, T., Karavritra, V., & Ahoniemi, T. (2009). Rubyric - An online assessment tool for effortless authoring of personalized feedback. In: *Proceedings of the Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE, Paris, France.

Bailey, R., & Szabo, Z. (2006). Assessing engineering design process knowledge. *International Journal Of Engineering Education*, *22(3),* 508-518.

Baxter, D. H., & Guerci, M. J. (2003, June). *Automating an introductory computer aided design course to improve student evaluation*. Presented at the 2003 ASEE Annual Conference & Exposition, Nashville, Tennessee.

Bertoline, G., Wiebe, E. N., Hartman, N. W., & Ross, W. A. (2011). *Fundamentals of graphics communication 6th ed*. New York, NY: McGraw-Hill.

Bhavnani, S. K., John, B. E., & Flemming, U. (1999). The strategic use of CAD: An empirically inspired theory-based course. In *CHI'99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp.183-90). New York: ACM Press.

Bodein, Y., Rose, B., & Caillaud, E. (2013). A roadmap for parametric CAD efficiency in the automotive industry. *Computer-Aided Design 2013, 45(10),* 1198–1214.

Bodein, Y., Rose, B., & Caillaud, E. (2014). Explicit reference modeling methodology in parametric CAD system. *Computers in Industry, 65(1),* 136-147.

Booker, J. D. (2012). A survey-based methodology for prioritising the industrial implementation qualities of design tools. *Journal of Engineering Design, 23(7),* 507-525.

Bracewell, R., Wallace, K., Moss, M., & Knott, D. (2009). Capturing design rationale. *Computer - Aided Design, 41(3),* 173-86.

Branoff, T. J., Hartman, N. W., & Wiebe, E. N. (2002). Constraint-based, three-dimensional solid modeling in an introductory engineering graphics course: re-examining the curriculum. *Engineering Design Graphics Journal, 66(1),* 5-10.

Branoff, T., & Wiebe, E. (2009). Face-to-face, hybrid, or online? Issues faculty face redesigning an introductory engineering graphics course. *Engineering Design Graphics Journal, 73(1)*, 25-31.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD books. Alexandria, Virginia, USA.

Brunnermeier, S.B., & Martin, S.A. (2002). Interoperability costs in the US automotive supply chain. *Supply Chain Management: An International Journal, 7(2)*, 71–82.

Burge, J., & Brown, D.C. (1998). *Design Rationale Types and Tools.* AI in Design Group. Technical Report, Worchester Polytechnic Institute, Computer Science Dept.

Cabrera, I., & Villalon, J. (2013). An adaptive interface for computer-assisted rubrics in an e-Marking tool using nearest neighbor. In: *Proceedings of the IEEE 13th International Conference on Advanced Learning Technologies*, 72-76.

Camba J., Contero M., & Company P. (2016). Parametric CAD modeling: An analysis of strategies for design reusability. *Computer-Aided Design, 74*, 18-31.

Camba J., Contero M., Johnson M., & Company P. (2014). Extended 3D annotations as a new mechanism to explicitly communicate geometric design intent and increase CAD model reusability. *Computer-Aided Design, 57*, 61-73

Camba J., Contero M., Otey J., & Company P. (2014, August). Explicit Communication of Geometric Design Intent in CAD: Evaluating Annotated Models in the Context of Reusability. *Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference DETC/DTM 2014.* August 17-20, 2014, Buffalo, New York, US. 26th International Conference on Design Theory and Methodology (DTM) Paper DETC2014-34527.

Camba, J., Contero, M., & Salvador-Herranz, G. (2014). Speak with the annotator: Promoting interaction in a knowledge-based CAD environment built on the extended annontatoin concept, *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design* (CSCWD), Hsinchu, pp. 196-201.

Camba, J., Cosin, A., & Contero, M. (2014, November). *An evaluation of formal strategies to create stable and reusable parametric feature-based 3D models.* In ASME International Mechanical Engineering Congress and Exposition (pp. V011T14A003-V011T14A003). American Society of Mechanical Engineers.

Carberry A. R., & McKenna A. F. (2014) Exploring student conceptions of modeling and modeling uses in engineering design. *Journal of Engineering Education, 103(1)*, 77–91.

Cebrián-Robles, D., Serrano-Angulo, J., & Cebrián-de-la-Serna, M. (2014). Federated eRubric service to facilitate self-regulated learning in the European university model. *European Educational Research Journal, 13(5),* 575-584.

Chandrasegaran, S. K., Ramani, K., Sriram, R. D., Horváth, I., Bernard, A., Harik, R. F., & Gao, W. (2013). The evolution, challenges, and future of knowledge representation in product design systems. *Computer-aided Design, 45(2),* 204-228.

Chang, K. H. (2014). *Product design modeling using CAD/CAE*. Amsterdam: Elsevier.

Chen, X., & Hoffman, C. M. (1995). *Design compilation of feature-based and constraint-based CAD*. Proceedings of the third ACM Symposium on Solid Modeling and Applications (pp. 13-19). ACM.

Chester, I. (2008). 3D-CAD: modern technology–outdated pedagogy? *Design and Technology Education: an International Journal, 12(1)*, 7-9.

Chirone, E., & Tornincasa, S. (2011). *Disegno tecnico industriale*. Torino: Edizioni Il Capitello. (In Italian)

Choi, G. H., Mun, D., & Han, S. (2002). Exchange of CAD part models based on the macro-parametric approach. *International Journal of CAD/CAM, 2(1),* 13-21.

Company, P., Contero, M., & Salvador-Herranz, G. (2013). *Testing Rubrics for Assessment of Quality in CAD Modelling REES 2013.* Research in Engineering Education Symposium 2013 Proceedings – Kuala Lumpur, ISBN 978-967-0194-35-6, pp 107-112.

Company P., Contero M., Otey J., Camba J.D., Agost M.J., &Pérez-López D.C. (2017). Web-based system for adaptable rubrics: case study on CAD assessment. *Journal of Educational Technology & Society*, 20(3), 24-41.

Company P., Contero M., Otey J., & Plumed R. (2015). Approach for developing coordinated rubrics to convey quality criteria in MCAD training. *Computer-Aided Design, 63,* 101–117.

Company, P., Gomez-Fabra, M., Agost M. J., & Vergara M. (2007). Assessment strategy to engage students in constraining parametric CAD drawings. *International Journal of Engineering Education, 23 (5),* 990-999.

Company, P., & González, C. (2013). 3D CAD with SolidWokrs. Ed. UJI (in Spanish). (http://cad3dconsolidworks.uji.es).

Company, P., Otey, J., Camba, J., & Contero, M. (2014, August). Leveraging mechanical 3D CAD systems through improved model quality based on best practices and rubrics. *Proceedings of the ASME 2014 International Design Engineering IDETC/CIE 2014*. Buffalo, New York, US.

Company P., Otey J., Contero M., Agost M.J., & Almiñana A. (2016). A. Implementation of Adaptable Rubrics for CAD Model Quality Formative Assessment. *International Journal of Engineering Education* (ISSN 0949-149X), 32(2A), 749–761.

Condoor, S. S. (1999, November). *Integrating design in engineering graphics courses using feature-based, parametric solid modeling*. Presented at the 29th ASEE/IEEE Frontiers in Education Conference (pp. 12d2-13), San Juan, Puerto Rico.

Contero, M., Company, P., Vila, C., & Aleixos N. (2002). Product Data Quality and Collaborative engineering. *IEEE Computer Graphics and Applications (ISSN 0272-1716), 22(3)*, 32-42.

Creighton, R. H. (2010). *Unity 3D game development by example: A seat-of-your-pants manual for building fun, groovy little games quickly*. Packt Publishing Ltd.

Crismond D. P., & Adams R. S. (2012) The informed design teaching and learning matrix. *Journal of Engineering Education, 101(4),* 738–797.

Dankwort, C. W., Weidlich, R., Guenther, B., & Blaurock, J. E. (2004). Engineers' CAx education—it's not only CAD. *Computer-Aided Design, 36(14)*, 1439–1450.

Dassault Systemes. (2015, July 31). *SolidWorks Help: Design Intent*. Retrieved from http://help.solidworks.com/2014/English/solidworks/acadhelp/c_design_intent.htm

Devine, K. L., & Laingen, M. A. (2013, October). *Assessing design intent in an introductory-level engineering graphics course.* Presented at the 68th Mid-Year Conference of the ASEE Engineering Design Graphics Division (pp. 59-63), Worcester, Massachusetts.

Diefes-Dux, H. A., Zawojewski, J. S., & Hjalmarson, M. A. (2010). Using educational research in the design of evaluation tools for open-ended problems. *International Journal of Engineering Education*, *26 (*4), 807-819

Dimopoulos, I., Petropoulos, O., & Retalis, S. (2013). Assessing students' performance using the learning analytics enriched rubrics. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 195-199.

Diwakaran, R. P., & Johnson M. D. (2012). Analyzing the effect of alternative goals and model attributes on CAD model creation and alteration. *Computer-Aided Design, 44(4),* 343-353.

Dorribo-Camba, J., & Contero, M. (2014, May). *Integration of model-based design annotations in product lifecycle management systems to facilitate design intent communication*. Presented at the 10th International Symposium on Tools and Methods of Competitive Engineering (pp. 1-10), Budapest, Hungary.

Economides, A. A., & Roupas, C. (2007). Evaluation of computer adaptive testing systems. *International Journal of Web Web-Based Learning and Teaching Technologies, 2(1)*, 70-87.

Felder, R. M., & Brent, R (2003). Designing and teaching courses to satisfy the ABET engineering criteria. *Journal of Engineering Education. 92(1),* 7–25.

Field D. A., (2004). Education and training for CAD in the auto industry. *Computer-Aided Design, 36*, 1431–1437.

Fischer, G., Lemke, A., McCall, R., & Morch, A. I. (1991). Making argumentation serve design. *Human-Computer Interaction, 6*, 393-419.

Freiesleben, J. (2010). Proposing a new approach to discussing economic effects of design quality. *International Journal of Production Economics, 124,* 348-359.

Fu, M. W., Ong, S. K., Lu, W. F., Lee, I. B. H., & Nee, A. Y. C. (2002). An approach to identify design and manufacturing features from a data exchanged part model. *Computer-Aided Design, 35(11),* 979-993.

Gebhard, R. (2013). A resilient modeling strategy [Internet]. *SolidEdge University 2013; [cited 2013 Aug 4]*. http://www.solidedgeu.com/presentations-2013.

Georgiadou, E., Triantafillou, E., & Economides, A. A. (2006). Evaluation parameters for computer-adaptive testing. *British Journal of Educational Technology, 37(2),* 261–278.

Gerbino, S. (2003). *Tools for the interoperability among CAD systems*. XIII ADM—XV INGEGRAF International Conference on Tools and Methods Evolution in Engineering Design.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57(4),* 2333-2351.

González-Lluch, C., Company, P., Contero, M., Camba, J. D., & Plumed, R. (2017). A survey on 3D CAD model quality assurance and testing tools. *Computer-Aided Design, 83,* 64-79.

González-Lluch, C., Company, P., Contero, M., Camba, J. D., & Colom, J. (2017). A case study on the use of model quality testing tools for the assessment of MCAD models and drawings. *International Journal of Engineering Education*. In press.

Goodrich, A. H. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership, 57(5),* 13–18.

Goodrich , A. H.(1996). Understanding rubrics. *Educational Leadership, 54,* 14-17.

Gu, H., Chase, T. R., Cheney, D. C., Bailey, T., & Johnson, D. (2001). Identifying, correcting, and avoiding errors in computer-aided design models which affect interoperability. *Journal of Computing and Information Science in Engineering, 1(2)*, 156-166.

Gwet, K. L. (2014). *Handbook of inter-rater reliability* (4th ed.). Gaithersburg, MD: Advanced Analytics, LLC.

Hamade, R. F., & Artail. H. A. (2010). A study of the influence of learning style of beginner computer-aided design users on their performance. *Journal of Engineering Design, 21(5),* 561-577

Hamade, R. F., & Artail, H. A. (2008). A study of the influence of technical attributes of beginner CAD users on their performance. *Computer-Aided Design, 40(2),* 262–272.

Hamade, R. F., Artail, H. A., & Jaber, M. Y. (2007). Evaluating the learning process of mechanical CAD students. *Computers and Education, 49(3),* 640–61.

Hamade, R. F., Artail, H. A., & Jaber, M. Y. (2005). Learning theory as applied to mechanical CAD training of novices. *International Journal of Human-Computer Interaction, 19(3),* 305–22.

Han, J. H. (1996). *Survey of feature research*. Technical Report IRIS-96-346. Institute for Robotics and Intelligent Systems: University of Southern California.

Han, J. H. (1997). On multiple interpretations. In C. Hoffmann and W. Bronsvort. (Eds.), *Proceedings of the Fourth Symposium on Solid Modeling and Applications* (pp. 311-321). ACM Press.

Harchay, A., Chniti-Belcadhi, L. A., & Braham, R. (2014). Context-Aware framework to provide personalized mobile assessment. *Interaction Design and Architecture(s) Journal - IxD&A, 23*, 82-97.

Hartman, N. W. (2004). The development of expertise in the use of constraint-based CAD tools: examining practicing professionals. *Engineering Design Graphics Journal, 68(2),* 4-26.

Hartman, N. W. (2005). Defining expertise in the use of constraint-based CAD tools by examining practicing professionals. *Engineering Design Graphics Journal, 69(1)*, 6-15.

Henderson, M. R., (1993, May). *Representing functionality and design intent in product models*. Proceedings on the second ACM Symposium on Solid Modeling and Applications (pp. 387-396). Montreal, Canada.

Horwood, M., & Kulkarni, S. (2005). CAD data quality. *Engineering Designer, 31(3),* 14–16.

Irwin, J. (2013). Wind turbine blade CAD models used as scaffolding technique to teach design engineers. *Engineering Design Graphics Journal, 77(2),* 1-17.

Isbell, T., & Goomas, D. T. (2014). Computer-Assisted rubric evaluation: Enhancing outcomes and assessment quality. *Community College Journal of Research and Practice, 38*, 1193–1197.

ISO. (2005a). ISO 9000:2005, *Quality management systems: Fundamentals and vocabulary*. Organisation Internationale de Normalisation. Geneve, Suisse.

ISO. (2005b). ISO 10303-55, *Industrial automation systems-Product data representation and exchange-Part 55: Integrated generic resource: Procedural and hybrid representation.* Organisation Internationale de Normalisation. Geneve, Suisse.

ISO. (2005c). ISO 10303-108, *Industrial automation systems and integration-Product data representation and exchange-Part 108: Integrated application resource: Parameterization and constraints for explicit geometric product models.* Organisation Internationale de Normalisation. Geneve, Suisse.

Iyer, G. R., & Mills, J. J. (2006). Design Intent in 2D CAD: Definition and Survey. *ComputerAided Design and Applications, 3(1-4),* 259-267.

Jackson, C., & Buxton, M. (2007). *The design reuse benchmark report. Seizing the opportunity to shorten product development*. Boston: Aberdeen Group.

Jackson, C., & Prawel, D. (2013). *The 2013 state of 3D collaboration and interoperability report (p. 7),* Lifestyle Insights and Longview Advisors.

Jarczyk, A. P., Loffler, P., & Shipman, F. M. (1992). Design rationale for software engineering: A survey. *Proceedings of the 25th Hawaii International Conference on System Sciences: Vol. 2. (pp. 577-586).* IEEE.

Johnson, M. D., & Diwakaran R. P. (2009, August-September). Assessing the effect of incentive on computer-aided design intent. *Proceedings of the ASME 2009 International Design Engineering Technical Conference & Computers and Information in Engineering Conference (pp.1-10)*. San Diego, CA.

Johnson, M. D., & Diwakaran R. P. (2011). An educational exercise examining the role of model attributes on the creation and alteration of CAD models. *Computers & Education 57,* 1749–1761.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*, 130-144.

Karkehabadi, S. (2013). Using Rubrics to Measure and Enhance Student Performance. Office of Institutional Research, Planning, and Assessment Northern Virginia Community College. www.nvcc.edu/

Karsenty, L. (1996, April). An empirical evaluation of design rationale documents. *Proceedings of the ACM CHI 96 Annual Conference on Computer-Human Interaction. Vancouver, BC Canada (pp. 150-156).*

Kimura, F., & Suzuki H. (1989). A CAD system for efficient product design based on design intent. *CIRP Annals-Manufacturing Technology. 38(1*), 149-152.

Kirstukas, S. J. (2016, June). *Development and evaluation of a computer program to assess student CAD models.* Proceedings of ASEE's 123rd Annual Conference and Exposition, New Orleans, LA.

Kunz, W., & Rittel, H. (1970). *Issues as Elements of Information Systems (Vol. 131).* Berkeley, California: Institute of Urban and Regional Development, University of California.

Landers, D. M., & Khurana, P. (2004). *US Patent No. 6,775,581*. Washington, DC: U. S. Patent and Trademark Office.

Laursen, E., & Møller, L. (2015). Collaboration between Industrial Designers and Design Engineers–Comparing the Understanding of Design Intent. *Studies in Health Technology and Informatics, 217,* 367.

Lawton, D., Vye N., Bransford, J., Sanders E., Richey, M., French, D., & Stephens, R. (2012). Online learning based on essential concepts and formative assessment. *Journal of Engineering Education, 101(2),* 244–287.

Leahy, K. (2012, November). Promoting best practice design intent in 3D CAD modelling: feedback as a constructivist paradigm of teaching and learning. *Proceedings of the ASEE EDGD 67th Midyear Meeting (pp. 59-66)*, Limerick, Ireland.

Lee, J. (1997), Design Rationale Systems: Understanding the Issues. *IEEE Expert, 12(3),* 78-85.

Lee, J., & K. Lai. (1991). What's in design rationale? *Human–Computer Interaction, 6(3-4),* 251-280.

Li, M., Langbein, F. C., & Martin, R. R. (2010). Detecting design intent in approximate CAD models using symmetry. *Computer-Aided Design, 42(3),* 183-201.

Lieblich, J. H. in association with B.R. Fischer. (2008). *Drawing requirements manual: guide for digital models, digital and traditional drawings, and technical data packages for commercial and military applications.* Global Engineering Documents. Englewood, CO.

Lieu, D., & Sorby, S. (2008). *Visualization, modeling, and graphics for engineering design 1st ed*. Boston, MA:Delmar Cengage Learning.

Liu, O. L., Brew, C., Blackmore, J., Gerard, L., Madhok, J., & Linn, M. C. (2014). Automated Scoring of Constructed-Response Science Items. *Prospects and Obstacles, Educational Measurement: Issues and Practice, 33(2)*, 19-28.

MacLean, A., Young, R. M., Bellotti, V. M., & Moran, T. P. (1991). Questions, options, and criteria: Elements of design space analysis. *Human–Computer Interaction, 6(3-4),* 201-250.

MacLean, A., Young, R. M., & Moran, T. P. (1989, April-June). Design rationale: the argument behind the artifact. In Bice, K. and Lewis, C. H. (Eds.), *Proceedings of the ACM CHI 89 Human Factors in Computing Systems Conference (pp. 247-252).* Austin, Texas.

Malini, R. Y., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35(4),* 435-448.

Mandorli, F., & Otto, H. E. (2013). Negative knowledge and a novel approach to support MCAD education. *Computer-Aided Design and Applications 2013, 10(6),* 1007–20.

Mandorli, F., Otto, H.E., & Raffaeli, R. (2015). Explicit 3D Functional Dimensioning to Support Design Intent Representation and Robust Model Alteration. *CAD'15 Proceedings*, pp 203-209.

Mandorli, F., Otto, H. E., & Raffaeli, R. (2016). Explicit 3D functional dimensioning to support design intent representation and robust model alteration. *Computer-Aided Design and Applications, 13(1),* 108-123.

Manson, J. R., & Olsen, R. J. (2010). Diagnostics and rubrics for assessing learning across the computational science curriculum. *Journal of computational Science 1*, 55-61.

Mathews, K, Janicki, T., He, L., & Patterson, L. (2012). Implementation of an automated grading system with an adaptive learning component to affect student feedback and response time. *Journal of Information Systems Education, 23(1),* 71-83.

Matuk, C. F., Linn, M. C., & Eylon, B. S. (2015). Technology to support teachers using evidence from student work to customize technology-enhanced inquiry units. *Instructional Science, 43(2),* 229-257.

McCormick, M., Lawyer, K., Wiggins, J., Swan, C., Paterson, K., and Bielefeldt, A. (2015). Sustainable engineering assessment using rubric-based analysis of challenge question responses. *Journal of Professional Issues in Engineering Education and Practice, 141(2).*

McKenney, D. (1998). Model quality: The key to CAD/CAM/CAE interoperability. In: *Proceedings of the 1998 MSC Software Americas Users Conference*. Universal City, CA.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom, Practical *Assessment, Research & Evaluation, 7(25)*, http://pareonline.net/getvn.asp?v=7&n=25

Miller, C. A., Funk, H., Goldman, R., Meisner, J., & Wu, P. (2005). Implications of adaptive vs. adaptable UIs on decision making: Why "Automated Adaptiveness" is not always the right answer. *Proceedings of the 1st International Conference on Augmented Cognition*, Las Vegas, NV; July 22-27.

Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching, 50(1),* 34–40.

Moodle. Online Learning Management system. Accessed May, 2015. https://docs.moodle.org/28/en/Features

Moskal, B. (2000). Scoring Rubrics: What, When and How? Practical Assessment, *Research & Evaluation, 7(3),* http://pareonline.net/getvn.asp?v=7&n=3

Mostow, J. (1985). Toward better models of the design process. *AI magazine. 6(1),* 44-57.

Mun, D., Han, S., Kim, J., & Oh, Y. (2003). A set of standard modeling commands for the history-based parametric approach. *Computer-Aided Design, 35(13),* 1171-1179.

Nakahara, T., Soga, T., Nakamura, Y., Mitani, M., & Kawana, N. (2014). Development of an e-Textbook connecting with learning management system and a study of its effective use. *EdMedia 2014, Association for the Advancement of Computing in Education*, Tampere, Finland: June 23-26, Paper #41578.

Nash, M. (2015). Computer adaptive assessments with Google forms. *Tech 2 Teach*. http://tech-2-teach.com/?p=66

Nield, B. S., Guzowski, R., Nassif, N., Simpson, A. M. & Martiniello-Wilks, R. (2014). First use of Re:View: A tool to combine assessment tasks, marking criteria and graduate attributes for biochemistry students. *International Journal of Innovation in Science and Mathematics Education*, *22(7),* 49-64.

Nitko, A. (2001). *Educational assessment of students* (3rd ed.). Upper Saddle River, Nj: Merrill.

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21(3),* 239–250.

Otey, J., Company, P., Contero, M., & Camba, J. D. (2014, June). A review of the design intent concept in the context of CAD model quality metrics. *ASEE Annual Conference and Exposition,* Indianapolis, IN. http://www.asee.org/conferences-and-events/conferences/annual-conference/2014. Paper ID #9556.

Otey J., Company P., Contero M., & Camba J.D. (2017). Revisiting the Design Intent Concept in the Context of Mechanical CAD Education. *Computer-Aided Design and Applications*,15(1), 1-14.

Otto, H. E., & Mandorli, F. (2015). A Framework to Support 3D Explicit Modeling Education and Practice. *Computer-Aided Design and Applications, 12(1)*, 104-117.

Owensby, J. E., & Summers, J. D. (2014). Assembly Time Estimation: Assembly Mate Based Structural Complexity Metric Predictive Modeling. *Journal of Computing and Information Science in Engineering, 14(1)*, 011004.

Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education, 54*, 715–721.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. H. (2007). *Engineering design: a systematic approach 3rd ed.* London: Springer-Verlag.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144.

Pena-Mora, F., Sriram, D., & Logcher, R. (1993). SHARED-DRIMS: SHARED Design Recommendation-Intent Management System. *Enabling Technologies: Infrastructure for Collaborative Enterprises, Proceedings, Second Workshop (pp.213-221).*

Piegl, L. A. (2005). Ten challenges in computer-aided design. *Computer-Aided Design, 37(4),* 461–470.

Plumed, R., Varley, P., & Company, P. (2012). Features and design intent in engineering sketches. In D. Plemenos, & G. Miaoulis (Eds.), *Intelligent Computer Graphics (pp. 77-106)*. Berlin Heidelberg: Springer-Verlag.

Pophan, W. J. (1997). What´s wrong and what´s right with rubrics?. *Educational Leadership, 55(2)*, 72-75.

Pratt, M. J., Anderson, B. D., & Ranger, T. (2005). Towards the standardized exchange of parameterized feature-based CAD models. *Computer-Aided Design, 37(12*), 1251-1265.

PTC Creo. (2015, July 31). *Designing with intent*. Retrieved from http://creo.ptc.com/2010/12/02/designing-with-intent/

Ramos-Barbero, B., & Garcia-Garcia, R. (2009). Strategic learning of simulation and functional analysis in the design and assembly of mechanisms with CAD on a professional master's degree course. *Computer Applications in Engineering Education, 19(1),* 146-160.

Ramos-Barbero, B., Melgosa-Pedrosa, C., & Zamora-Samperio, R. (2016). Learning CAD at university through summaries of the rules of design intent. *International Journal of Technology and Design Education,* 1-18.

Rayón, A., Guenaga, M., & Nuñez A. (2014). Supporting competency-assessment through a learning analytics approach using enriched rubrics. In: *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM '14)*, 291-298.

Regli, W. C., Hu, X., Atwood, M., & Sun, W. (2000). A survey of design rationale systems: approaches, representation, capture and retrieval. *Engineering with Computers, 16(3-4),* 209-235.

Reid, K. J., & Cooney, E. M. (2008). Implementing rubrics as part of an assessment plan. *Internationa Journal of Engineering Education, 24(5),* 893-900.

Rejab, M. M., Hassan, S., Awang, I., & Ahmad, M. (2011). Designing customizable rubrics tool to support formative evaluation of problem-based learning courses. In: *Proceedings of the International Conference on Computer Technology and Development,* 3rd (ICCTD 2011). ASME Press.

Requicha, A. A. G. (1980). Representations for rigid bodies: theory, methods, and systems. *ACM Computing Surveys. 12(4),* 437-464.

Requicha, A. A. G., & Rossignac, J.R. (1992). Solid modeling and beyond. *IEEE Computer Graphics and Communications, 12(5)*, 31-44.

Resilient Modeling Strategy™ V4. (2017, May 13). retrieved from http://learnrms.com.

Rohrmann, B. (2007). Verbal qualifiers for rating scales: Sociolinguistic considerations and psychometric data. *Project Report. University of Melbourne, Australia*.

Rossignac, J. (2004). Education-driven research in CAD. *Computer-Aided Design, 36(14),* 1461-1469.

Rynne, A., & Gaughran, W. (2008). Cognitive modeling strategies for optimum design intent in parametric modelling. *Computers in Education Journal. 18(1*), 55-68.

Santos, P., Colina X., Hernández-Leo, D., Melero, J., & Blat, J. (2009). Enhancing Computer Assisted Assessment Using Rubrics in a QTI Editor. In: *Proceedings of the IEEE International Conference on Advanced Learning Technologies*, ICALT, 303-305.

Sapidis, N. S., & Kim, M-S. (2004). Editorial to special issue: CAD education. *Computer-Aided Design, 36(14),* 1429–1430.

Shih, C. H., & Anderson, B. (1997). A design/constraint model to capture design intent. *Proceedings of the fourth ACM Symposium on Solid Modeling and Applications (pp. 255-264).* ACM.

Shum, S., & Hammond, N. (1993), *Argumentation-Based Design Rationale: From Conceptual Roots to Current Use.* Tech. Report EPC-1993-106, Rank Xerox Research Centre, Cambridge.

Siemens. (2015, July 31). *Designing products in NX.* Retrieved from http://www.plm.automation.siemens.com/en_us/products/nx/

Silva, J. F., & Restivo, F. J. (2012). An adaptive assessment system with knowledge representation and visual feedback. In: *Proceedings of the 15th International Conference on Interactive Collaborative Learning* (ICL) 1-4.

Sim, S., & Duffy, A. (1994), A New Perspective to Design Intent and design Rationale, *Artificial Intelligence in Design Workshop Notes for Representing and Using Design Rationale*, *1518 August,* 4-12.

Smit, R., & Birri, T. (2014). Assuring the quality of Standards-oriented classroom assessment with rubrics for complex competencies. *Studies in Educational Evaluation 43*, 5-13.

Srinivasan, V., Chakrabarti, A., & Lindemann, U. (2012). A framework for describing functions in design. *Proceedings of the 12th International Design Conference (pp. 1111-1121).*

Strategic Automotive product data Standards Industry Group, SASIG, Product Data Quality Workgroup (PDQ) (2005). SASIG– Product Data Quality Guidelines for the Global Automotive Industry, Guideline version 2.1 (STEP Part 59. ISO/PAS 26183:2006. URL: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43436)

Szykman, S., Sriam, R. D., & Regli, W. C. (2001). The role of knowledge in next-generation product development systems. *Journal of Computing and Information Science in Engineering, 1(1),* 3-11.

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation, 9(2)*, http://pareonline.net/getvn.asp?v=9&n=2

Toto, R., Colledge, T., Frederick, D., & Pung, W. H. (2014). Instructional strategies to promote student strategic thinking when using SolidWorks. *Advances in Engineering Education, 4(1)*, 1-37.

Ullman, D. (2003). *The mechanical design process.* NY (USA): McGraw-Hill.

Ullman, D. G. (2002). Toward the ideal mechanical engineering design support system. *Research in Engineering Design, 13(2),* 55-64.

Vandenbrande, J. H., & Requicha, A. A. G. (1993). Spatial reasoning for the automatic recognition of machinable features in solid models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(12),* 1269-1285.

VDA, Association of the German Automotive Industry. (1999). Scope and Quality of CAD/CAM Data. VDA Recommendation 4955/1 Testing and controlling the data quality of CAD models.

Waks, S., Trotskovsky, E., Sabag, N., & Hazzan, O. (2011). Engineering thinking: The experts' perspective. *International Journal of Engineering Education, 27(4),* 838-851.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems 1996, 12(4),* 5–34.

Wang, Z. S., Tian, L., Wu, Y. H., & Liu, B. B. (2015). Personalized knowledge push system based on design intent and user interest. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science.*

Wiebe, E. N. (1999). Future applications of geometry and graphics. *Engineering Design Graphics Journal, 63(2)*, 13-20.

Will, P. M. (1991). Simulation and modeling in early concept design: an industrial perspective. *Research in Engineering Design, 3(1*), 1-13.

Yang, J., Han, S., Kang, H., & Kim, J. (2006). Product data quality assurance for e-manufacturing in the automotive industry. *International Journal of Computer Integrated Manufacturing, 19(2)*, 136– 147.

Ye, X., Peng, W., Chen, Z., & Cai, Y-Y. (2004). Today's students, tomorrow's engineers: An industrial perspective on CAD education. *Computer Aided Design, 36(14),* 1451-1460.

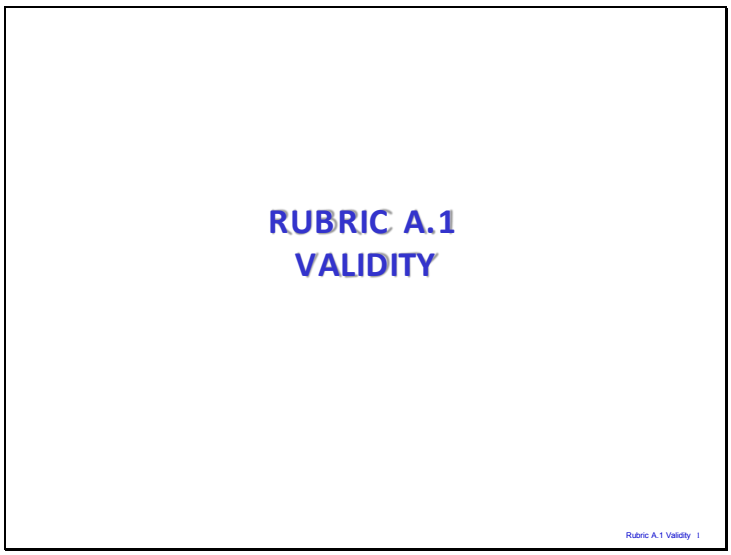
Zaiontz, C. (2015). Real Statistics Using Excel. Last access: June 2015. URL: http://www.real-statistics.com/

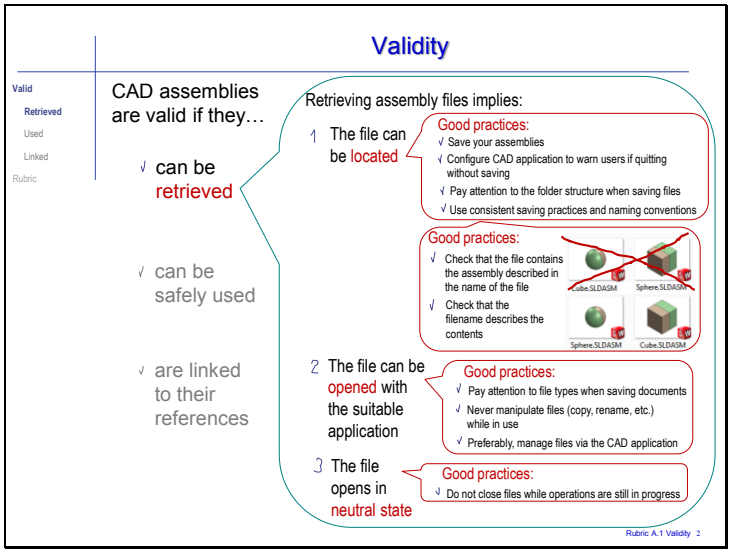
Zhang, Y., X., Luo, J., Li, & Buis, J. J. (2013). A semantic representation model for design rationale of products. *Advanced Engineering Informatics, 27(1),* 13-26.

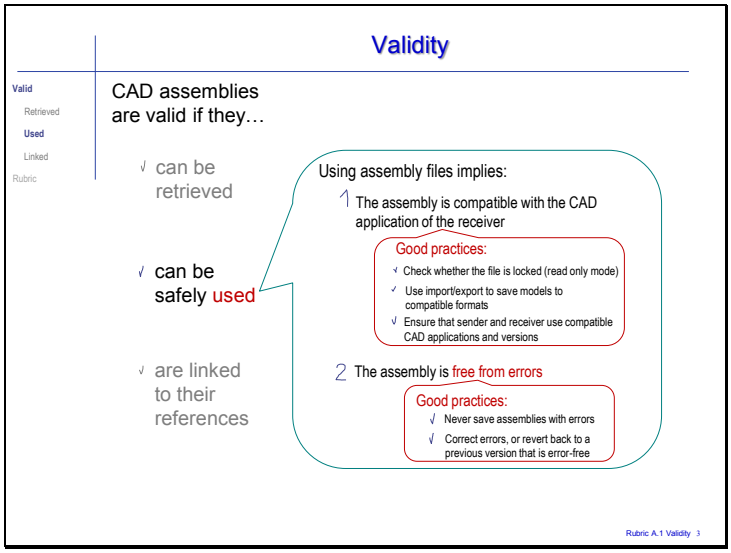
Zhang, Y., & Luo, X. (2009, March-April). Design intent information exchange of feature-based CAD models. *Proceedings of the World Congress on Computer Science and Information Engineering (pp. 11-15),* Los Angeles, CA.

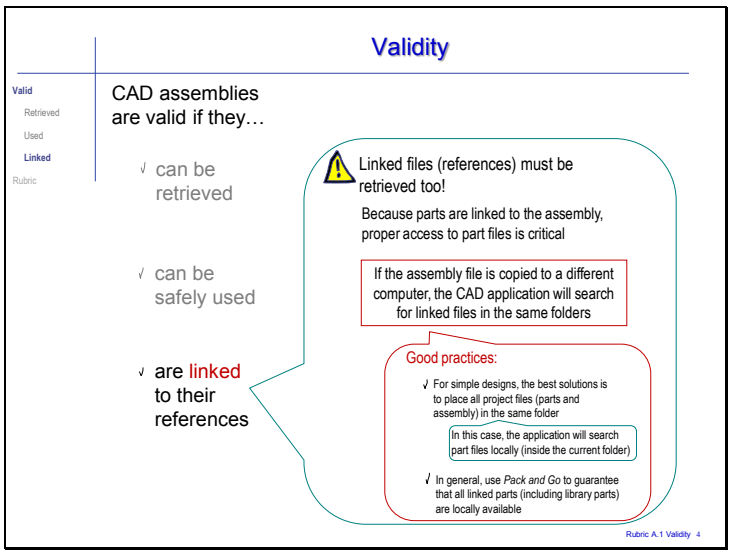
Appendix A
Assembly Rubric Presentations

All slides developed and created by Pedro Company and provided to illustrate methods utilzed to instruct students in the use of rubrics, for those who may wish to incorporate them into their classes.




Validity
Valid
Retrieved
Used
Linked
Rubric
CAD assemblies are valid if they…
√ can be retrieved
√ can be safely used
√ are linked to their references
Retrieving assembly files implies:
1 The file can be located
Good practices:
√ Save your assemblies
√ Configure CAD application to warn users if quitting without saving
√ Pay attention to the folder structure when saving files
√ Use consistent saving practices and naming conventions
Good practices:
√ Check that the file contains the assembly described in the name of the file
√ Check that the filename describes the contents
Cube.SLDASM
Sphere.SLDASM
Sphere.SLDASM
Cube.SLDASM
2 The file can be opened with the suitable application
Good practices:
√ Pay attention to file types when saving documents
√ Never manipulate files (copy, rename, etc.) while in use
√ Preferably, manage files via the CAD application
3 The file opens in neutral state
Good practices:
√ Do not close files while operations are still in progress
Rubric A.1 Validity 2


Validity
Valid
Retrieved
Used
Linked
Rubric
CAD assemblies are valid if they…
√ can be retrieved
√ can be safely used
√ are linked to their references
Using assembly files implies:
1 The assembly is compatible with the CAD application of the receiver
Good practices:
√ Check whether the file is locked (read only mode)
√ Use import/export to save models to compatible formats
√ Ensure that sender and receiver use compatible CAD applications and versions
2 The assembly is free from errors
Good practices:
√ Never save assemblies with errors
√ Correct errors, or revert back to a previous version that is error-free
Rubric A.1 Validity 3


Validity
Valid
Retrieved
Used
Linked
Rubric
CAD assemblies are valid if they…
√ can be retrieved
√ can be safely used
√ are linked to their references
Linked files (references) must be retrieved too!
Because parts are linked to the assembly, proper access to part files is critical
If the assembly file is copied to a different computer, the CAD application will search for linked files in the same folders
Good practices:
√ For simple designs, the best solutions is to place all project files (parts and assembly) in the same folder
In this case, the application will search part files locally (inside the current folder)
√ In general, use Pack and Go to guarantee that all linked parts (including library parts) are locally available
Rubric A.1 Validity 4

Rubric

Valid
Retrieved
Used
Linked
Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
A1.1	The file of the assembly can be located and opens in neutral state
A1.1a	The file of the assembly has the expected contents (and name) and is in the expected folder (or website)
A1.1b	The file of the assembly can be re-opened after closing the current session (even on a different computer)
A1.1c	The file of the assembly opens in neutral state (i.e. no operations are in progress)
A1.2	The assembly can be used
A1.2a	The assembly is compatible with the CAD of the receiver
A1.2b	The assembly tree is free from error messages
A1.3	All components linked to the assembly may be accessed (including parts, sub-assemblies and library parts), even when libraries are not available, or when software compatibility issues exist between versions
A1.3a	All parts linked to the assembly can be accessed
A1.3b	All sub-assemblies linked to the assembly can be accessed
A1.3c	All library parts linked to the assembly can be accessed

Rubric A.1 Validity 5

RUBRICS A.2 COMPLETENESS

Rubric A.1 Validity 6

Completeness: parts

Complete
Parts
Standard parts
Placement
Rubric

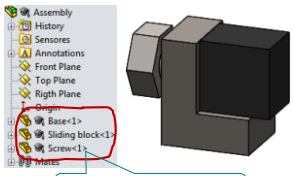
An assembly model is complete if it…

1 Includes all the necessary components

2 Uses standard parts when required

3 All parts are correctly placed

Use the assembly tree to check that all parts have been included in the assembly model



The number of instances of each part must also match!

But see next page!

Good practices:
- Check that all parts in the assembly are included in the assembly tree

Rubric A.1 Validity 7

Completeness: parts

Complete
- Parts
  - Standard parts
  - Placement

Rubric

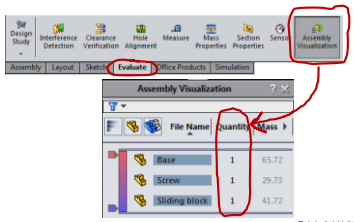
There is a problem with the instance counter in SolidWorks

When reinserting parts after being deleted, the Instance ID's will numerically skip the deleted items

This means that the ID counter is unreliable

The only solution to force a reset of the ID counters is:
1. Eliminate the out-of-sequence instances of the part
2. Save the assembly
3. Quit and re-start SolidWorks
4. Re-insert the parts

The best way to check the number of copies of each part in an assembly is by using *Assembly Visualization*



Rubric A.1 Validity 8

Completeness: parts

Complete
- Parts
  - Standard parts
  - Placement

Rubric

An assembly model is complete if it…

1. Includes all the necessary components
2. Uses standard parts when required
3. All parts are correctly placed

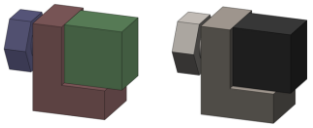
It is good practice to use different colors for different parts in an assembly

There are two different criteria to select colors:

Contrast different parts ↔ Increase realism

Use high contrast colors, if possible

Use colors that resemble the actual colors of the materials



Select *contrast* to check completeness of the assembly and ease detection of other assembly failures

Rubric A.1 Validity 9

Completeness: parts

Complete
- Parts
  - Standard parts
  - Placement

Rubric

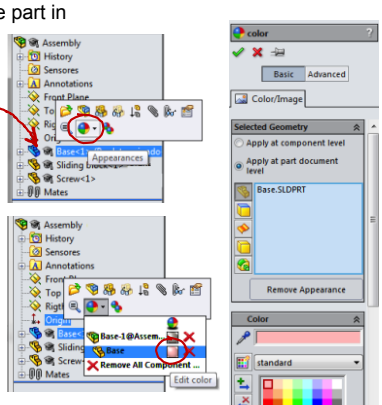
To assign color to one part in an assembly:

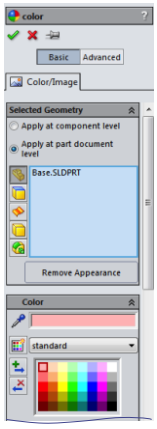
√ Select the part in the assembly tree

√ Select *Appearances* in the contextual menu of the part

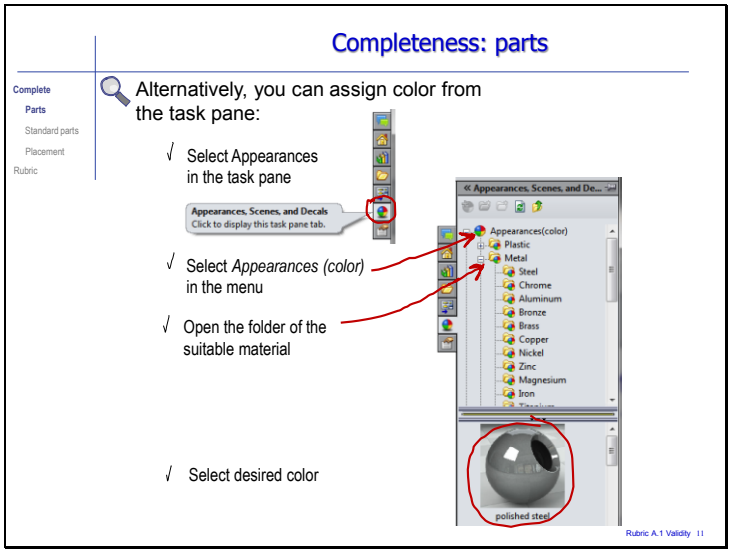
√ Click *Edit color*

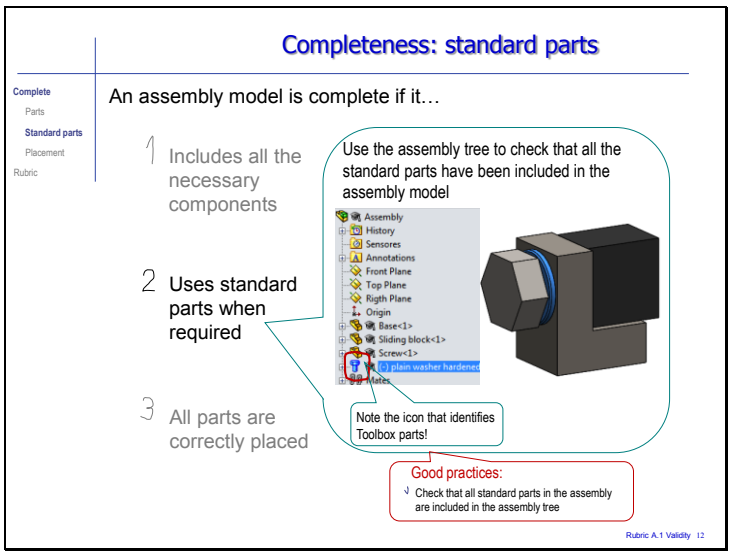
√ Select the appropriate color

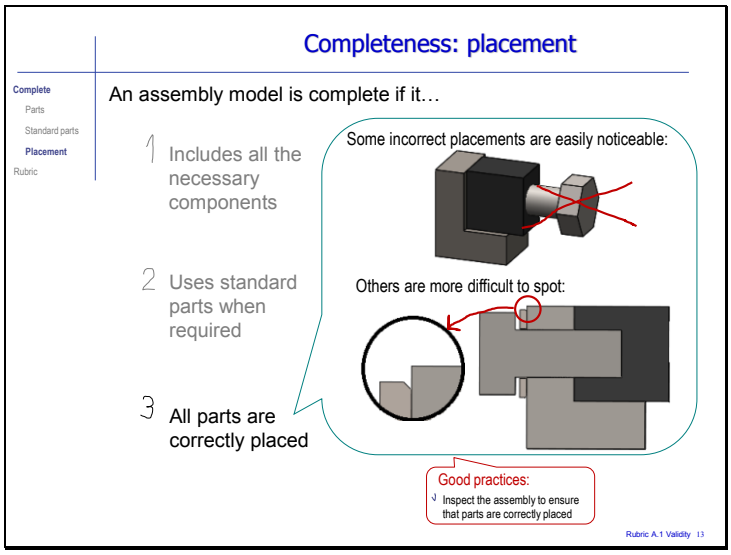




Rubric A.1 Validity 10


Completeness: parts
Complete
Parts
Standard parts
Placement
Rubric
Alternatively, you can assign color from the task pane:
√ Select Appearances in the task pane
√ Select Appearances (color) in the menu
√ Open the folder of the suitable material
√ Select desired color
Appearances, Scenes, and Decals
Click to display this task pane tab.
Appearances(color)
Plastic
Metal
Steel
Chrome
Aluminum
Bronze
Brass
Copper
Nickel
Zinc
Magnesium
Iron
polished steel
Rubric A.1 Validity 11


Completeness: standard parts
Complete
Parts
Standard parts
Placement
Rubric
An assembly model is complete if it…
1 Includes all the necessary components
2 Uses standard parts when required
3 All parts are correctly placed
Use the assembly tree to check that all the standard parts have been included in the assembly model
Assembly
History
Sensores
Annotations
Front Plane
Top Plane
Rigth Plane
Origin
Base<1>
Sliding block<1>
Screw<1>
Mates
Note the icon that identifies Toolbox parts!
Good practices:
Check that all standard parts in the assembly are included in the assembly tree
Rubric A.1 Validity 12


Completeness: placement
Complete
Parts
Standard parts
Placement
Rubric
An assembly model is complete if it…
1 Includes all the necessary components
2 Uses standard parts when required
3 All parts are correctly placed
Some incorrect placements are easily noticeable:
Others are more difficult to spot:
Good practices:
Inspect the assembly to ensure that parts are correctly placed
Rubric A.1 Validity 13

Completeness: placement

Complete
Parts
Standard parts
Placement
Rubric

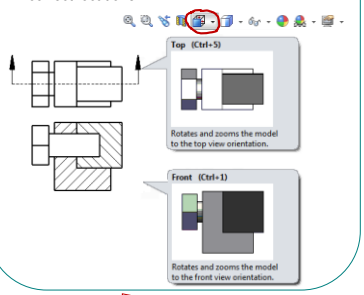
An assembly model is complete if it…

1 Includes all the necessary components

2 Uses standard parts when required

3 All parts are correctly placed

A systematic search may be helpful to inspect the model and check that all parts are in the correct locations



Good practices:
- Use views to check whether parts are correctly placed

Completeness: placement

Complete
Parts
Standard parts
Placement
Rubric

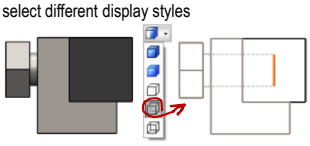
An assembly model is complete if it…

1 Includes all the necessary components

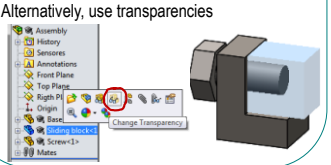
2 Uses standard parts when required

3 All parts are correctly placed

For internal parts, use the *Show* menu to select different display styles



Alternatively, use transparencies



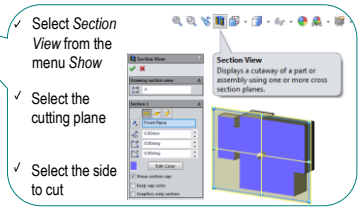
Good practices:
- Use all view styles to facilitate part inspection

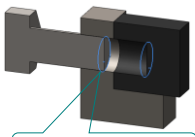
Completeness: placement

Complete
Parts
Standard parts
Placement
Rubric

Use *Section views* to ease the assembly process

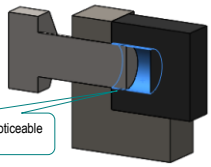
- ✓ Select *Section View* from the menu *Show*
- ✓ Select the cutting plane
- ✓ Select the side to cut





The cut eases the selection of internal elements for adding mates

Section Views can also be used to check whether parts have been correctly placed



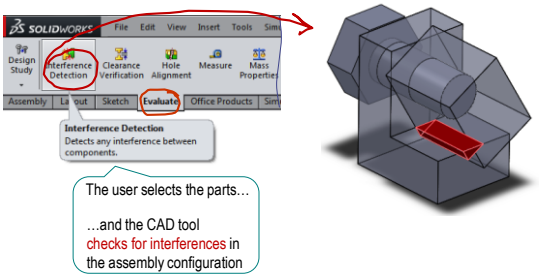
Incorrect placement, only noticeable after creating section view!

Completeness: placement

Complete
- Parts
- Standard parts
- Placement

Rubric

Assembly modelers include tools to check if components are incorrectly placed:



The user selects the parts…

…and the CAD tool checks for interferences in the assembly configuration

Rubric

Complete
- Parts
- Standard parts
- Placement

Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
A2.1	The assembly includes all and only necessary parts and sub-assemblies
A2.1a	The assembly includes all the parts and sub-assemblies in the assembly
A2.1b	The assembly includes all necessary copies of each part and sub-assembly in the assembly
A2.1c	The assembly is free from surplus and alien parts and sub-assemblies
A2.2	Standard library parts are included when required, which are suitably instantiated from the library
A2.2a	Standard library parts are used when required
A2.2b	Standard library parts are suitably instantiated rom the library
A2.3	Components (parts, sub-assemblies and library parts) are correctly placed and free of interferences
A2.3a	Relative location among components (parts, sub-assemblies and library parts) match their functional positions
A2.3b	Components (parts, sub-assemblies and library parts) are free of unwanted interferences

RUBRICS A.3 CONSISTENCY

Consistency

Consistent
Rubric

For simple products, the assembly model must allow designers to perform analysis

→ An assembly model can only be successfully used for analysis if it is consistent

In addition to being valid and complete!

For complex designs, assembly models are also the primary view of the digital representation of new industrial products

→ Secondary views are derived for other purposes (like digital mock-up, analysis, or manufacturing)

The primary view must be consistent to allow for such secondary views!

Rubric A.1 Validity 20

Consistency

Consistent
Base part
Constraints
Rubric

An assembly model is consistent if:

1. Assembly is correctly placed and linked

The assembly must be:
- √ Upright
- √ Centered
- √ Symmetrically placed

This is important because for some types of analyses, the assembly must interact with its environment

It is also important when the assembly becomes a sub-assembly that must fit into a larger assembly

2. Assembly allows valid motion and prevents undesired motion

The assembly must allow for CAE analysis (like kinematics for mechanisms, or structural stability for rigid objects)

Rubric A.1 Validity 21

Consistency

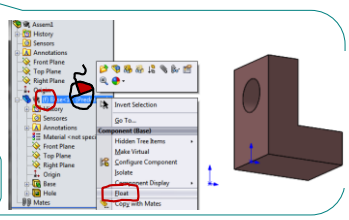
Consistent
Base part
Constraints
Rubric

1. Since all parts are linked to the base part, the placement of the assembly in the scene depends on the placement of the base part

In fact, CAE modules (like tolerance analysis) usually assume that the base part is fixed

Good practices:
- √ Link the base part to the global reference system

- √ If the base part is fixed by default, make it *float*
- √ Make the origin of the part coincident with the origin of the global system



Alternatively, select the homonym planes and use mates to make them coincident

Rubric A.1 Validity 22

Consistency

Consistent
Base part
Constraints
Rubric

2 In general, all components must be suitably assembled by way of mate conditions

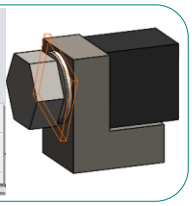
Good practices:
Check the prefixes in components names in the Feature Manager design tree:
No prefix= fully defined
(?) = not solved
(+) = over defined
(-) = under defined
(f) = fixed

Some type motion is irrelevant to analyze the behavior of the assembly, therefore it is not necessary to constrain it

For instance, it usually makes no sense to limit the rotation of a washer

Therefore, it acceptable to let some parts under constrained

Although you can lock some of them easily when desired!



Rubric A.1 Validity 23

Consistency

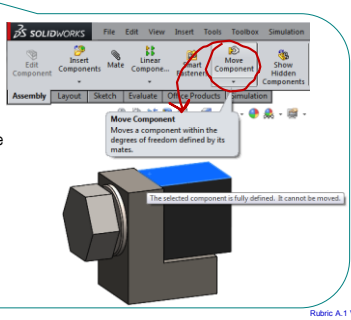
Consistent
Base part
Constraints
Rubric

Some assemblies are mechanisms, thus they should not be fully constrained

Good practices:
Try moving the different parts of the assembly to check whether they move only the way they should

User selects one part…

…and tries to interactively move it by dragging it with the cursor



Rubric A.1 Validity 24

Rubric

Consistent
Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
A3.1	The assembly is placed and linked relative to the global reference system
A3.1a	The only fixed part is the base part, or assemblies with multiple fixed components have proper base
A3.1b	The base component is correctly linked to the global reference system
A3.2	All components are suitably assembled by way of mate conditions (assembly allows valid motion and prevents undesired motion)
A3.2a	Assembly prevents invalid motion
A3.2b	Assembly allows valid motion (removal of minimum degrees of freedom necessary to function)

Rubric A.1 Validity 25

RUBRICS A.4
CONCISENESS

Conciseness

Concise
Rubric

CAD assemblies are concise if they:

- √ Use suitable mates
  - Do not contain repetitive of fragmented mating conditions
- √ Use high semantics
  - Use high-level assembly operations when available
- √ Are short linked
  - Do not contain unnecessarily long chains of linked parts

Conciseness: mates

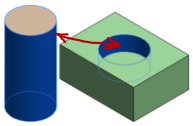
Concise
Mates
Patterns
Links
Rubric

Analyze mates to check that the assembly:

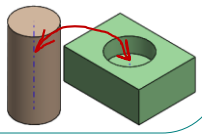
- √ Does not contain repetitive mates
- √ Does not contain fragmented mates

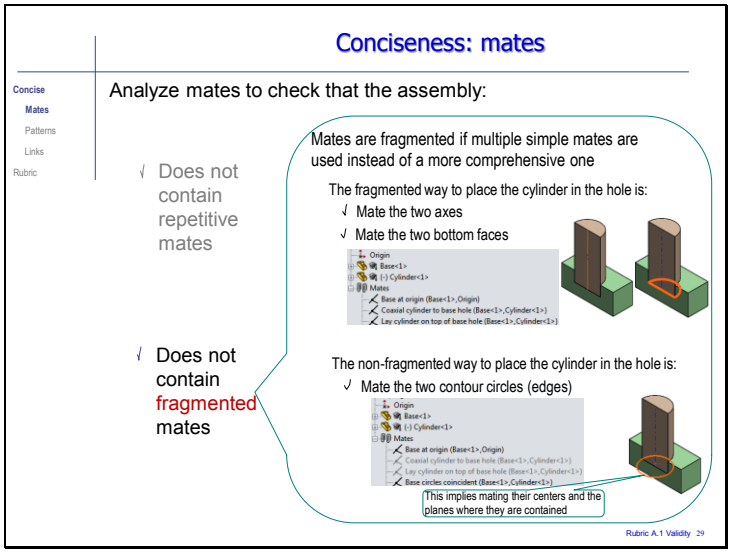
Mates are repetitive if they fully or partly re-constrain the same Degrees of Freedom

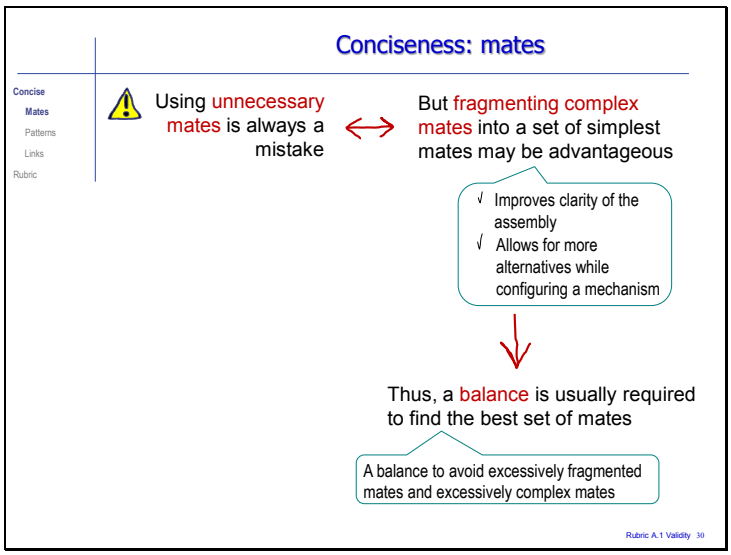
If one cylinder is concentric with a hole…

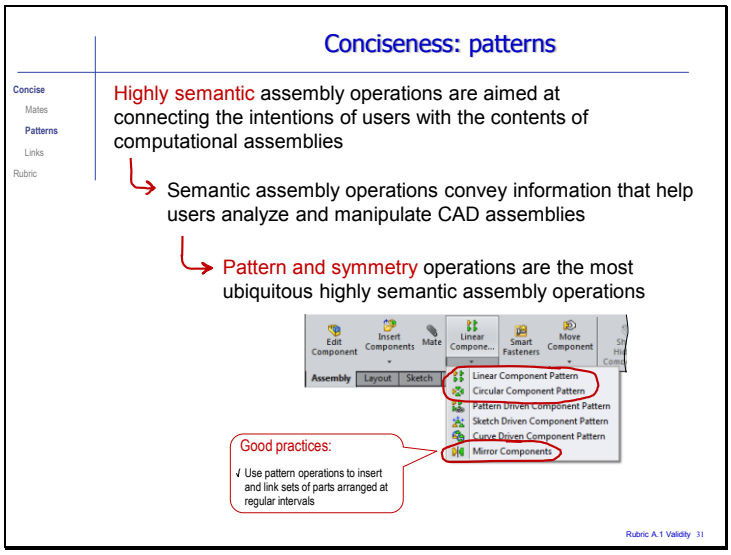


…it is repetitive to mate both as coaxial




Conciseness: mates
Concise
Mates
Patterns
Links
Rubric
Analyze mates to check that the assembly:
√ Does not contain repetitive mates
√ Does not contain fragmented mates
Mates are fragmented if multiple simple mates are used instead of a more comprehensive one
The fragmented way to place the cylinder in the hole is:
√ Mate the two axes
√ Mate the two bottom faces
The non-fragmented way to place the cylinder in the hole is:
√ Mate the two contour circles (edges)
This implies mating their centers and the planes where they are contained
Rubric A.1 Validity 29


Conciseness: mates
Concise
Mates
Patterns
Links
Rubric
Using unnecessary mates is always a mistake
But fragmenting complex mates into a set of simplest mates may be advantageous
√ Improves clarity of the assembly
√ Allows for more alternatives while configuring a mechanism
Thus, a balance is usually required to find the best set of mates
A balance to avoid excessively fragmented mates and excessively complex mates
Rubric A.1 Validity 30

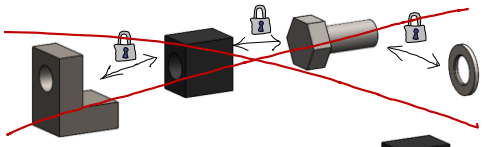

Conciseness: patterns
Concise
Mates
Patterns
Links
Rubric
Highly semantic assembly operations are aimed at connecting the intentions of users with the contents of computational assemblies
Semantic assembly operations convey information that help users analyze and manipulate CAD assemblies
Pattern and symmetry operations are the most ubiquitous highly semantic assembly operations
Edit Component
Insert Components
Mate
Linear Compone...
Smart Fasteners
Move Component
Assembly
Layout
Sketch
Linear Component Pattern
Circular Component Pattern
Pattern Driven Component Pattern
Sketch Driven Component Pattern
Curve Driven Component Pattern
Mirror Components
Good practices:
√ Use pattern operations to insert and link sets of parts arranged at regular intervals
Rubric A.1 Validity 31

Conciseness: links

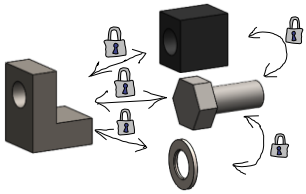
Concise
Mates
Patterns
Links
Rubric

It is not good practice to define long chains of relations among different parts, as it may result in unforeseen relations!

Also, calculation times increase significantly and the application becomes more prone to round off and similar errors



It is better practice to use a short subset of base parts and relate the remaining parts directly to them!



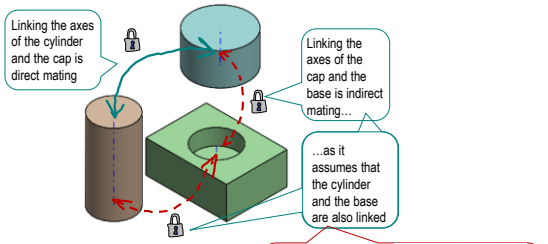
Rubric A.1 Validity 32

Conciseness: links

Concise
Mates
Patterns
Links
Rubric

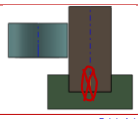
Indirect mating is bad practice since:

- ✗ Makes the mating task more difficult
- ✗ Prevents editing mates to re-arrange the assembly



Good practices:

- ✓ Ensure that you select the correct axes, particularly when overlapping axes exist!



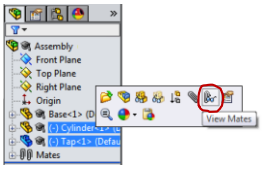
Rubric A.1 Validity 33

Conciseness: links

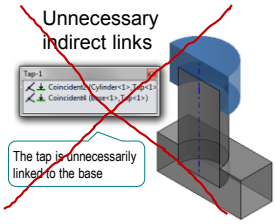
Concise
Mates
Patterns
Links
Rubric

Use *View Mates* tool to detect long chains and indirect links:

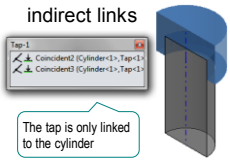
- ✓ Select a component in the assembly tree
- ✓ Click View Mates in the contextual menu
- ✓ Examine the parts linked to the selected parts



Unnecessary indirect links



Free of indirect links



Rubric A.1 Validity 34

Rubric

Concise
Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
A4.1	The assembly is free from repetitive or fragmented mating conditions
A4.2	Replication operations (translate-and-repeat, rotate-and-repeat and symmetry) are used whenever possible
A4.2a	Pattern operations (translate-and-repeat, rotate-and-repeat) are used whenever possible
A4.2b	Symmetry (if exists) is used to define the assembly
A4.3	The parent/child relations in the assembly tree are free of unnecessary dependences

Rubric A.1 Validity 35

RUBRICS A.5 CLARITY

Rubric A.1 Validity 36

Clarity

Clarity
Rubric

Since CAD assemblies are documents shared by different stakeholders during the design process, *communication* is important

In order to facilitate communication, the document:

- Must be clear and understandable (aimed at being understood at first glance)
  1. Mating operations must be labeled in the assembly tree to emphasize their function, instead of the type of link
  2. Related mating operations must be grouped in the assembly tree to emphasize their relationships
- Must follow conventions
  3. Most compatible mating operations are always preferred
  4. Standard mating operations are always preferred

Rubric A.1 Validity 37

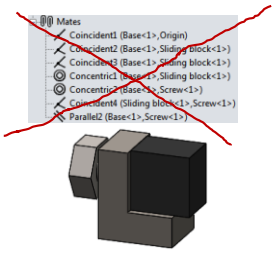
Clarity

Clarity
Rubric

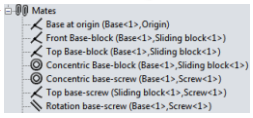
1 Mating operations are automatically labeled in the assembly tree

→ But the system only knows *how* the operations were created

→ However, function (*what for*) is much more important



Recommendation:
√ Re-label the mates to emphasize what they link, instead of the type of link

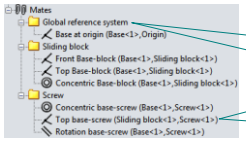


Clarity

Clarity
Rubric

2 Mating operations may be grouped according to different valid grouping criteria:

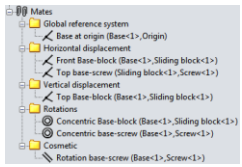
√ Grouping by parts



To create a new folder, right click and select *Add to New Folder* in the contextual menu

Drag and drop to add one mate to an existing folder

√ Grouping by Degrees of Freedom



Absolute "best grouping" does not exist!

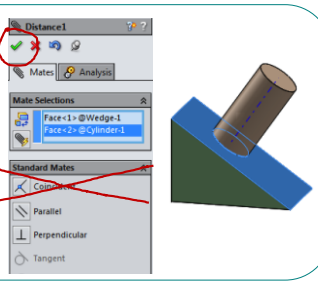
Recommendation:
√ Instead of searching for the "best" solution, simply avoid clearly bad solutions

Clarity

Clarity
Rubric

3 When different mates are available, select the simplest and most compatible

For instance, making the base of the cylinder coincident to the oblique face of the wedge is simpler than putting them at a distance of zero

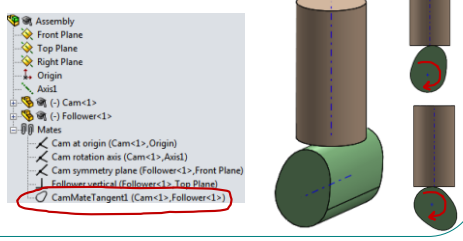


Clarity

Clarity
Rubric

4 Use high level mates when they are standard or common

For instance, *cam* is suitable for a particular type of mechanisms:



Agreements are required, as using sophisticated and exclusive mates may reduce the portability of a CAD assembly

Rubric A.1 Validity 41

Rubric

Clarity
Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
M5.1	All components, sub-assemblies, and mate constraints are properly labeled and organized in folders
M5.1a	Components are labeled and grouped to emphasize their function, instead of how they were defined
M5.1b	Mates are labeled to emphasize their function, instead of how they were defined
M5.1c	Related mates are grouped to emphasize parent-child relationships
M5.2	The assembly uses compatible and standard mates
M5.2a	The most compatible mates are always used
M5.2b	The most standard mates are always used

Rubric A.1 Validity 42

RUBRICS A.6 DESIGN INTENT

Rubric A.1 Validity 43

Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

Most design methodologies use assembly models to analyze the behavior of the design

Different aspects can be analyzed:

1. Assembly planning — Where the assembly sequence is the main concern
2. Assembly Process Design (APD) — Where the functionalities are considered
3. Design for assembly (DFA) — Where the affordances of the parts to assemble and disassemble are analyzed
4. Varieties — Which considers entire product families instead of isolated products

Assembly models are said to convey Design Intent if they contain explicit information required for such analyses

Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

Assembly planning is the process of determining a set of instructions for mechanically assembling a product from a group of components

Task Planning is an additional translation which is usually needed to map these instructions onto robot operations

→ An assembly algorithm specifies assembly, disassembly, maintenance, or repair operations and their order

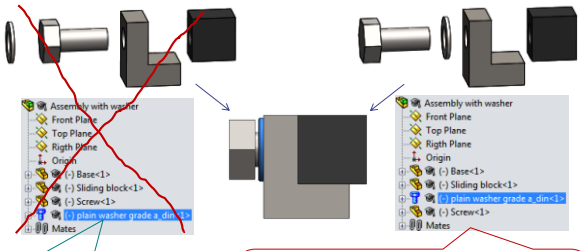
→ Sequencing is the core of assembly planning

→ To assist the assembly planning, the assembly sequence depicted in the assembly tree must replicate the assembly/disassembly process

Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

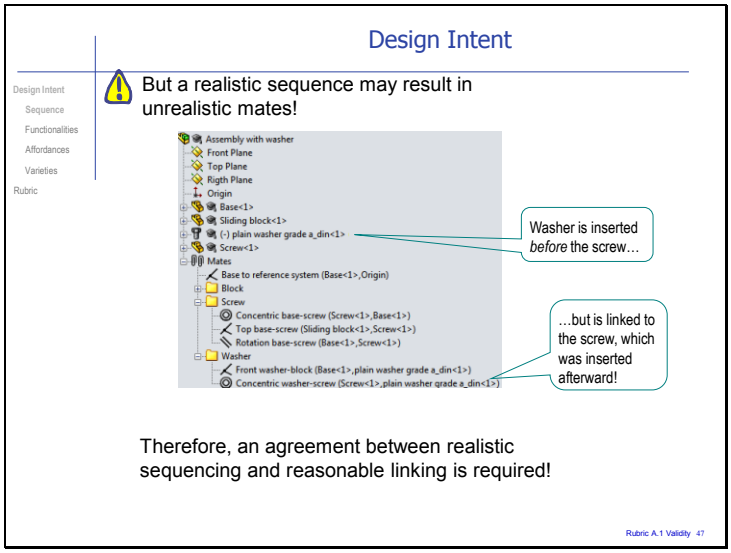
Inspect the assembly tree to determine if the assembly sequence mimics a realistic assembly sequence

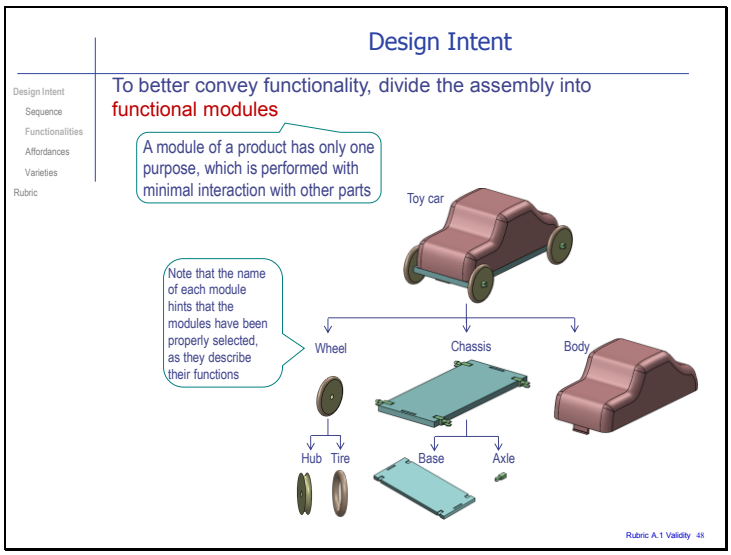


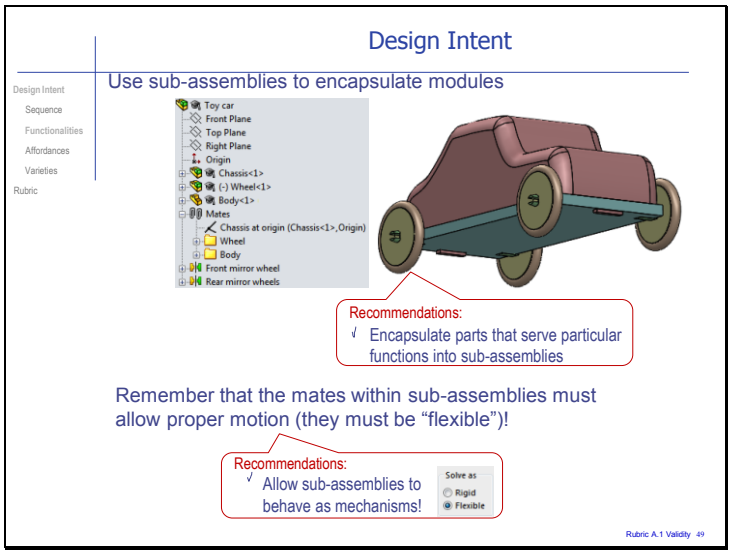
Drag and drop the components in the assembly tree, if necessary, to define the right sequence

Recommendations:

- √ The assembly sequence must go from main to auxiliary elements
- √ The disassembly sequence must be clearly described by reversing the assembly tree


Design Intent
Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric
But a realistic sequence may result in unrealistic mates!
Assembly with washer
Front Plane
Top Plane
Rigth Plane
Origin
Base<1>
Sliding block<1>
(-) plain washer grade a_din<1>
Screw<1>
Mates
Base to reference system (Base<1>,Origin)
Block
Screw
Concentric base-screw (Screw<1>,Base<1>)
Top base-screw (Sliding block<1>,Screw<1>)
Rotation base-screw (Base<1>,Screw<1>)
Washer
Front washer-block (Base<1>,plain washer grade a_din<1>)
Concentric washer-screw (Screw<1>,plain washer grade a_din<1>)
Washer is inserted *before* the screw…
…but is linked to the screw, which was inserted afterward!
Therefore, an agreement between realistic sequencing and reasonable linking is required!
Rubric A.1 Validity 47


Design Intent
Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric
To better convey functionality, divide the assembly into functional modules
A module of a product has only one purpose, which is performed with minimal interaction with other parts
Toy car
Note that the name of each module hints that the modules have been properly selected, as they describe their functions
Wheel
Chassis
Body
Hub
Tire
Base
Axle
Rubric A.1 Validity 48


Design Intent
Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric
Use sub-assemblies to encapsulate modules
Toy car
Front Plane
Top Plane
Right Plane
Origin
Chassis<1>
(-) Wheel<1>
Body<1>
Mates
Chassis at origin (Chassis<1>,Origin)
Wheel
Body
Front mirror wheel
Rear mirror wheels
Recommendations:
√ Encapsulate parts that serve particular functions into sub-assemblies
Remember that the mates within sub-assemblies must allow proper motion (they must be "flexible")!
Recommendations:
√ Allow sub-assemblies to behave as mechanisms!
Solve as
Rigid
Flexible
Rubric A.1 Validity 49

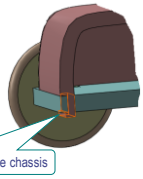
Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

Design for Assembly is a methodology in which parts are designed to include affordances that facilitate the assembly process

Affordances are features provided within parts to make them easier to grasp, move, orient and insert

Those affordances oriented at assembling are modeled as *Mating Features*



Fastener clip to fix the body to the chassis

Parts that include Mating Features should be assembled by taking advantage of their affordances

Features designed to facilitate grasping, moving, orienting, and inserting parts should be used to mate the parts while assembling

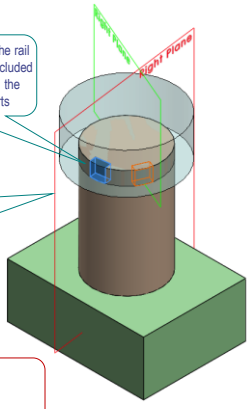
Rubric A.1 Validity 50

Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

Search for mating features before assembling…

The slot in the cylinder and the rail in the cap are affordances included by the designer to guarantee the correct alignment of both parts



…and use them to mate the parts

Since the features are aligned with their main reference systems, aligning the Right Planes of both parts would produce a geometrically equivalent assembly…

…but at the cost of ignoring the design intent conveyed by those affordances

Recommendation:
- Use mating features to assemble, whenever possible

Rubric A.1 Validity 51

Design Intent

Design Intent
Sequence
Functionalities
Affordances
Varieties
Rubric

Assembly Process Design (APD) handles product variety

We are interested in *functional variety*, which describes any differentiation in attributes related to the functionality of the product

There are two extreme approaches to manage product variety:

A process-based approach increases the flexibility of the assembly equipment

Increasing assembly equipment, increases manufacturing costs

This approach cannot be simulated with assembly modelers

A product-based approach standardizes most parts and offers diversity by varying the remaining parts

Component sharing or standardization reduces product diversity

Parts that should be replaced to allow for variety (product family), must be easily and safely replaced in the assembly model

Rubric A.1 Validity 52

Design Intent

Design Intent
- Sequence
- Functionalities
- Affordances
- Varieties

Rubric

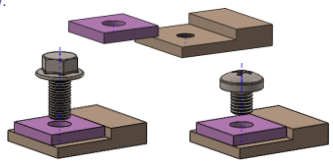
To provide support for variety in the assembly model, we must increase the independence of the parts that will be replaced:

If the part to replace is the screw:

√ Place the plate on top of the step, and mate their holes

√ Then, place the screw…

…which can be easily replaced

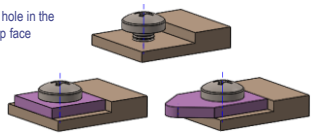


If the part to replace is the small plate:

√ Make the screw concentric with the hole in the step, and offset its head from the top face

√ Then, insert the smaller plate in between…

…so it is easy to replace



Rubric A.1 Validity 53

Rubric

Design Intent
- Sequence
- Functionalities
- Affordances
- Varieties

Rubric

The criteria described so far can be checked against an evaluation rubric

#	Criterion
A6.1	The assembly tree replicates the assembly/disassembly process
A6.1a	The assembly sequence proceeds from main to auxiliary elements
A6.1b	The assembly tree sequence reflects a realistic assembly sequence
A6.2	Sub-assemblies have been properly identified and efficiently used
A6.2a	Sub-assemblies encapsulate clearly perceived functions
A6.2b	The mate constraints of sub-assemblies provide for proper motion (they have been unfrozen)
A6.3	Mating features provided to ease assembly, if any, are mostly used for mating
A6.3a	Mating features provided to grasp, move, orient and insert the part, if any, have been identified
A6.3b	Mating features provided to grasp, move, orient and insert the part, if any, are mostly used for mating
A6.4	Parts that belong to modular families (if any) can be easily and safely replaced
A6.4a	Parts that belong to modular families have been identified
A6.4b	Parts that belong to modular families have been mated so as to be easily and safely replaced

Rubric A.1 Validity 54

Appendix B
Statistics Results

Wilcoxon Signed Rank Test:

S: Student

C: Instructor 1 (Company)

O: Instructor 2 (Otey)

Midterm Exam

Valid

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	23	266.00	0.000

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	22	242.50	0.000

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	3	2.00	0.789

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	0.125	(0, 0.25)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	0.125	(0, 0.25)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	0	(0, 0)	95.00%

Complete

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	-0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	29	72.50	0.002

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	-0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	30	45.00	0.000

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	13	18.00	0.059

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	-0.125	(-0.165, -0.04)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	-0.125	(-0.17, -0.045)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	0	(-0.04, 0)	95.00%

Consistent

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	-0.03

Test

Null hypothesis	H_0: $\eta = 0$
Alternative hypothesis	H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	33	219.00	0.276

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	-0.065

Test

Null hypothesis	H_0: $\eta = 0$
Alternative hypothesis	H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	33	191.00	0.112

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	0

Test

Null hypothesis H_0: $\eta = 0$

Alternative hypothesis H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	14	19.50	0.041

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	-0.03	(-0.125, 0.03)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	-0.065	(-0.125, 0)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	0	(-0.065, 0)	95.00%

Concise

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	0.045

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	38	488.00	0.090

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	0.04

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	38	423.50	0.446

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	-0.04

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	20	37.00	0.012

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	0.045	(0, 0.125)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	0.04	(-0.045, 0.085)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	-0.04	(-0.04, 0)	95.00%

Clear

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	-0.025

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	36	311.50	0.741

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	-0.065

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	37	263.00	0.184

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	-0.06

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	19	31.50	0.011

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	-0.025	(-0.125, 0.065)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	-0.065	(-0.13, 0)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	-0.06	(-0.065, 0)	95.00%

Design Intent

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	46	0.335

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	44	980.00	0.000

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	46	0.23

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	42	874.50	0.000

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	50	-0.085

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	37	9.00	0.000

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	46	0.335	(0.29, 0.415)	94.95%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	46	0.23	(0.17, 0.29)	94.95%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	50	-0.085	(-0.125, -0.08)	95.00%

Final Exam

Valid

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	16	68.00	1.000

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	16	68.00	1.000

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	0	(0, 0)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	0	(0, 0)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	0	(0, 0)	95.04%

Complete

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	-0.02

Test

Null hypothesis	H_0: $\eta = 0$
Alternative hypothesis	H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	34	227.00	0.231

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	-0.04

Test

Null hypothesis	H_0: $\eta = 0$
Alternative hypothesis	H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	36	206.00	0.047

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	13	0.00	0.002

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	-0.02	(-0.065, 0)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	-0.04	(-0.085, 0)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	0	(-0.04, 0)	95.04%

Consistent

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	0.065

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	41	606.00	0.023

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	43	503.00	0.722

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	51	-0.065

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	29	0.00	0.000

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	0.065	(0, 0.19)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	0	(-0.065, 0.065)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	-0.065	(-0.065, -0.065)	95.04%

Concise

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	-0.0625

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	44	393.50	0.239

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	-0.08

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	44	324.00	0.047

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	6	0.00	0.036

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	-0.0625	(-0.125, 0.04)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	-0.08	(-0.125, 0)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	0	(0, 0)	95.04%

Clear

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	-0.115

Test

Null hypothesis	H_0: $\eta = 0$
Alternative hypothesis	H_1: $\eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	40	207.50	0.007

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	-0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	40	149.00	0.000

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	51	0

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	8	0.00	0.014

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	-0.115	(-0.155, -0.02)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	-0.125	(-0.185, -0.06)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	0	(0, 0)	95.04%

Design Intent

Wilcoxon Signed Rank Test: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median
Diff S C	51	0.125

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S C	43	781.50	0.000

Wilcoxon Signed Rank Test: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median
Diff S O	51	0.04

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff S O	43	581.00	0.194

Wilcoxon Signed Rank Test: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median
Diff C O	51	-0.08

Test

Null hypothesis $H_0: \eta = 0$

Alternative hypothesis $H_1: \eta \neq 0$

Sample	N for Test	Wilcoxon Statistic	P-Value
Diff C O	39	0.00	0.000

Wilcoxon Signed Rank CI: Diff S C

Method

η: median of Diff S C

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S C	51	0.125	(0.065, 0.19)	95.04%

Wilcoxon Signed Rank CI: Diff S O

Method

η: median of Diff S O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff S O	51	0.04	(-0.02, 0.1)	95.04%

Wilcoxon Signed Rank CI: Diff C O

Method

η: median of Diff C O

Descriptive Statistics

Sample	N	Median	CI for η	Achieved Confidence
Diff C O	51	-0.08	(-0.125, -0.08)	95.04%

Pearson Correlation:

S: Student

C: Instructor 1 (Company)

O: Instructor 2 (Otey)

Midterm Exam

Valid

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.59945**	**1.**
	R Standard Error	0.01424	
	t	5.02398	
	p-value	8.49282E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.59945	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.55992**	**1.**
	R Standard Error	0.01526	
	t	4.53326	
	p-value	0.00004	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.55992	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.97734**	**1.**
	R Standard Error	0.00091	
	t	32.31887	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.97734	51	

Complete

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.59943**	**1.**
	R Standard Error	0.01424	
	t	5.02371	
	p-value	8.50038E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.59943	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.62977**	**1.**
	R Standard Error	0.01341	
	t	5.43864	
	p-value	2.10895E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.62977	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.93756**	**1.**
	R Standard Error	0.00247	
	t	18.86933	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.93756	51	

Consistent

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.49836**	**1.**
	R Standard Error	0.0167	
	t	3.85611	
	p-value	0.00036	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.49836	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.52457**	**1.**
	R Standard Error	0.01611	
	t	4.13326	
	p-value	0.00015	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.52457	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.95118**	**1.**
	R Standard Error	0.00194	
	t	21.57259	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.95118	51	

Concise

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.47598**	**1.**
	R Standard Error	0.01719	
	t	3.63066	
	p-value	0.00072	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.47598	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.50092**	**1.**
	R Standard Error	0.01665	
	t	3.88254	
	p-value	0.00034	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.50092	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.97104**	**1.**
	R Standard Error	0.00117	
	t	28.4487	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.97104	51	

Clear

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.36862**	**1.**
	R Standard Error	0.0192	
	t	2.66008	
	p-value	0.01079	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.36862	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.32739**	**1.**
	R Standard Error	0.01984	
	t	2.32432	
	p-value	0.02468	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.32739	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.9231**	**1.**
	R Standard Error	0.00302	
	t	16.80344	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.9231	51	

Design Intent

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.43827**	**1.**
	R Standard Error	0.01795	
	t	3.2709	
	p-value	0.00206	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.43827	47	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.60178**	**1.**
	R Standard Error	0.01417	
	t	5.05453	
	p-value	7.67002E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.60178	47	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.84657**	**1.**
	R Standard Error	0.00578	
	t	11.13339	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.84657	51	

Final Exam

Valid

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.73822**	**1.**
	R Standard Error	0.0091	
	t	7.73837	
	p-value	4.23081E-10	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.73822	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.73822**	**1.**
	R Standard Error	0.0091	
	t	7.73837	
	p-value	4.23081E-10	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.73822	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**1.**	**1.**
	R Standard Error	0.	
	t	2.14748E+10	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	1.	52	

Complete

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.73258**	**1.**
	R Standard Error	0.00927	
	t	7.61012	
	p-value	6.69564E-10	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.73258	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.73512**	**1.**
	R Standard Error	0.00919	
	t	7.66745	
	p-value	5.45299E-10	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.73512	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.98461**	**1.**
	R Standard Error	0.00061	
	t	39.83982	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.98461	52	

Consistent

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.31092**	**1.**
	R Standard Error	0.01807	
	t	2.31322	
	p-value	0.02486	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.31092	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.35272**	**1.**
	R Standard Error	0.01751	
	t	2.66545	
	p-value	0.01033	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.35272	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.97757**	**1.**
	R Standard Error	0.00089	
	t	32.82361	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.97757	52	

Concise

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.48536**	**1.**
	R Standard Error	0.01529	
	t	3.92541	
	p-value	0.00027	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.48536	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.5751**	**1.**
	R Standard Error	0.01339	
	t	4.9709	
	p-value	8.21851E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.5751	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.96389**	**1.**
	R Standard Error	0.00142	
	t	25.59472	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.96389	52	

Clear

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.54167**	**1.**
	R Standard Error	0.01413	
	t	4.55653	
	p-value	0.00003	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.54167	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.60329**	**1.**
	R Standard Error	0.01272	
	t	5.34892	
	p-value	2.20477E-6	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.60329	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.98004**	**1.**
	R Standard Error	0.00079	
	t	34.85951	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.98004	52	

Design Intent

S:C

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column E**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column E	**R**	**0.43169**	**1.**
	R Standard Error	0.01627	
	t	3.38403	
	p-value	0.0014	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column E vs. Column D	0.43169	52	

S:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column D	**Column F**
Column D	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.48355**	**1.**
	R Standard Error	0.01532	
	t	3.90626	
	p-value	0.00028	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column D	0.48355	52	

C:O

Correlation Coefficients Matrix			
Missing values removal	Pairwise deletion		
		Column E	**Column F**
Column E	**R**	**1.**	
	R Standard Error		
	t		
	p-value		
	H0 (5%)		
Column F	**R**	**0.95149**	**1.**
	R Standard Error	0.00189	
	t	21.86774	
	p-value	0.	
	H0 (5%)	*rejected*	
R			
Variable vs. Variable	*R*	*No# of valid cases*	
Column F vs. Column E	0.95149	52	

Appendix C

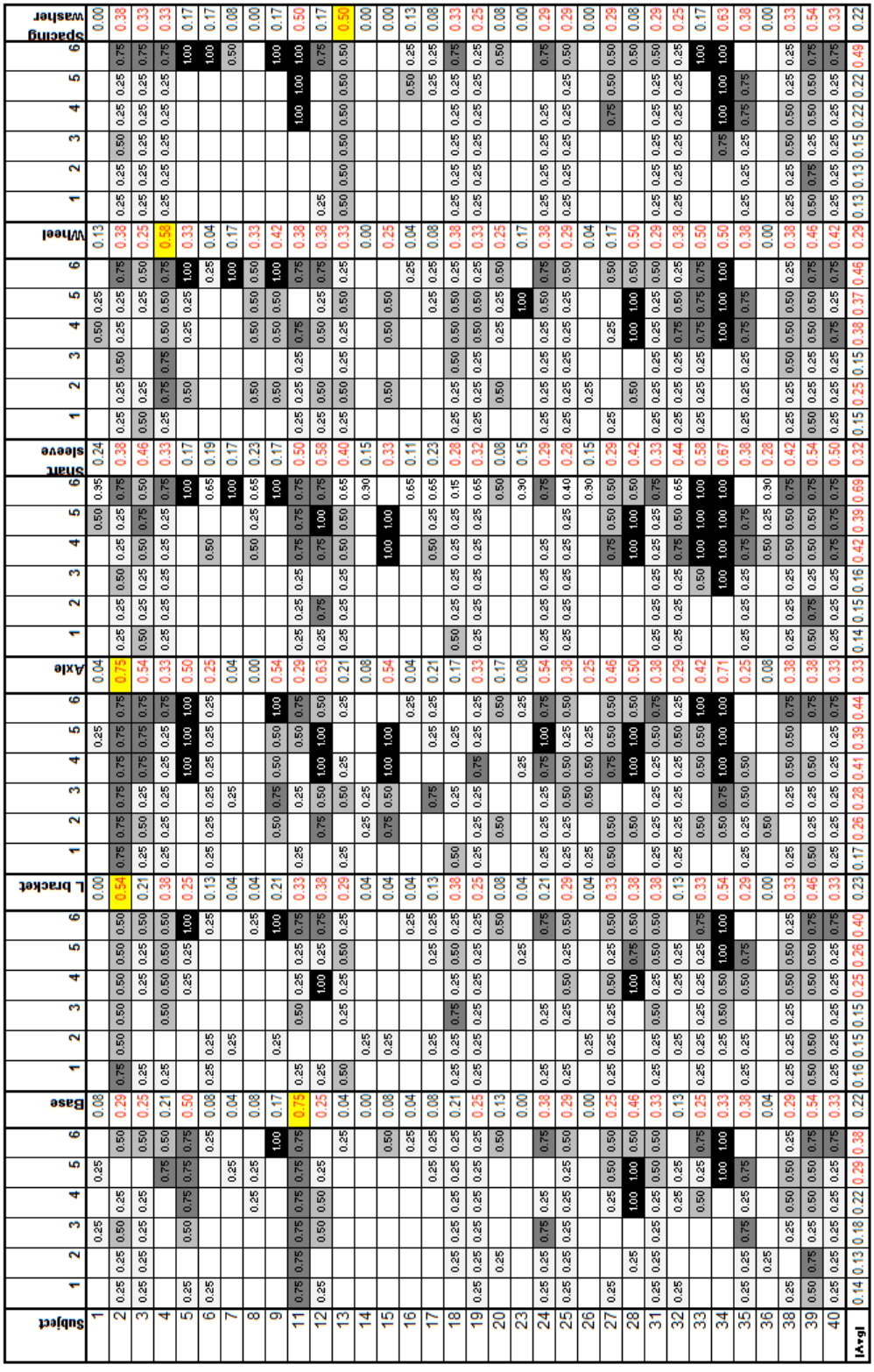



Table 3.7 Enlarged