Otey, J.; Company, P.; Contero, M.; Camba, JD. (2019). Assessment of parametric assembly models based on CAD quality dimensions. Computer-Aided Design and Applications. 16(4):628-653. https://doi.org/10.14733/cadaps.2019.628-653

Assessment of Parametric Assembly Models Based on CAD Quality Dimensions

Jeffrey Otey[0000-0002-3763-8759][1], Pedro Company[0000-0001-6399-4717][2], Manuel Contero[0000-0002-6081-9988][3] and Jorge D. Camba[0000-0001-5384-3253][4]

[1]Texas A&M University, j-otey@tamu.edu
[2]Universitat Jaume I, pcompany@uji.es
[3]Universitat Politecnica de Valencia, mcontero@upv.es
[4]Purdue University, jdorribo@purdue.edu

ABSTRACT

An approach to convey CAD quality-oriented strategies to beginning users to create bottom-up assemblies is described. The work builds on previous efforts in the area of single part history-based, feature-based parametric modeling evaluation by defining, testing, and validating a set of quality dimensions that can be applied to MCAD assembly assessment. The process of redefining and adapting dimension descriptors and achievement levels of parts rubrics to make them applicable to assemblies is addressed, then the results of two experimental studies designed to analyze the inter-rater reliability of this approach to assembly evaluation are reported. Results suggest the mechanism is reliable to provide an objective assessment of assembly models. Limitations for the formative self-evaluation of CAD assembly skills are also identified.

Keywords: CAD assembly, CAD quality, rubrics, MCAD education.

1 INTRODUCTION

The increasing importance of digital product models as core elements of new product development processes has been consistently supported by advances in the area of CAD data quality. Standards such as ISO 10303-59 [26] or the “Strategic Automotive product data Standards Industry Group” (SASIG) Product Data Quality (PDQ) Guideline V2.1 [49], for example, provide a set of basic criteria to ensure product data quality in diverse industrial settings. Other standards such as ASME Y14.41-2012 [3] and ISO 16792:2015 [27] regulate digital product definition data practices, which are notably relevant in Model-Based Engineering environments [34].

Digital product quality is concerned with identifying, locating, and solving data integrity problems in a master CAD model [35]. These problems may affect the simplification, interoperability, and (to a certain extent) reusability of the models [51],[22],[55]. In this paper, we focus on the master assembly model, its role in the overall quality of a digital product, and the significance of an effective assessment. Despite the extensive body of work in the field of Model Quality Testing (MQT) (summarized by Yang et al. [55], and more recently by González-Lluch et al. [19]) and the availability of commercial tools such as CADIQ, iCheck IT, and CT Core 3D Evolution, most MQT techniques and software mechanisms are limited to CAD quality errors of relatively low semantic level, i.e., they focus on the mathematical, geometric, and topological consistency and interoperability of CAD data. Many quality aspects of high-semantic level, e.g., those related to complexity, reusability, and the semantics associated to the modeling procedure such as design intent and design rationale considerations, have not received comparable attention [39],[13],[40].

Ignoring the high-semantic levels of CAD data quality can have negative impacts on model alteration and reusability, two fundamental benefits predicated by history-based, feature-based parametric CAD systems [5]. These aspects can, in turn, hinder the overall design reusability in the product development process by causing inefficiencies, errors, delays, and increasing cost [9],[28]. In this context, the quality of the master CAD model, i.e. the one created in the native CAD modeler (typically, a parametric feature-based, solid modeling MCAD system), is crucial, as this model serves as the primary source from which all secondary models used in CAx applications derive.

From an educational standpoint, CAD data quality (particularly at high-semantic levels) has major implications in training and student learning. Previous studies have shown that strategic knowledge (the ability of a user to identify proper procedures and select the most appropriate alternative to maximize the efficiency of a CAD tool) is essential to efficient CAD usage [5]. Some strategies have been suggested to teach this type of knowledge in an explicit manner [10],[47],[24] by providing interventions that actively encourage 3D CAD users to visualize, deconstruct, analyze, and cognitively assemble objects. As stated by Rynne and Gaughran [47], the ultimate goal of these interventions is to facilitate the development of “a mental model of parametric modeling systems in which the syntactic knowledge of the specifics of a system is supported by semantic knowledge of the tools available for creating and manipulating geometry in any system.”

In the domain of parametric solid modeling of parts, rubrics have been proven particularly effective for conveying and assessing high-semantic level quality criteria by way of good practices [14]. Furthermore, Company et al. [13] described an effective organization scheme of quality criteria in the form of rubrics that can used with parametric models of individual parts. The authors reported that rubrics with varying levels of detail are required to deliver the material at different phases along the CAD training period, a mechanism which we speculate can be equally effective when applied to assemblies. Nevertheless, the use of rubrics to assess assembly models and the interactions between its components, as well as the unique characteristics of assembly modeling techniques, remain relatively unexplored.

In this paper, we investigate the CAD quality dimensions of assembly models and study how these dimensions differ from those used in individual parts and how they must be organized, delivered, and assessed in order to successfully communicate high-semantic quality criteria in an explicit manner and from the early stages of instruction. The work builds on our previous studies on CAD quality assessment of single parametric models by providing a proactive approach to embed quality concepts for assembly modeling in the form of analytical rubrics that both communicate and evaluate CAD quality criteria. We present the results of two experiments in which the interrater reliability of our approach was evaluated as the agreement of assessment between instructors and between instructors and students.

2 RELATED WORK

Authors agree that most instructional strategies focus on declarative and procedural knowledge [11] but do not provide sufficient strategic knowledge or exposure to CAD quality concepts [5],[40].

Student feedback and evaluation are considered essential components of CAD instruction. Methods such as comparisons against a master CAD model, deviations from an agreed-upon “correct solution,” and written descriptions addressing specific mistakes or inefficient modeling approaches are often used to assess the knowledge and skills acquired in CAD. Research suggests that feedback should be well-timed, non-graded, and treated as a natural part of the learning process, so students can incorporate best practices and understand the importance of design methods and standards when using CAD [32],[42]. However, proper assessment of CAD models and assemblies is difficult and time-consuming, particularly in large classroom environments [8]. Some attempts to automate the assessment process have been reported. A review of common CAD assessment approaches is shown in Table 1.

Year	Authors	Solution	Domain
2003	Baxter and Guerci [3]	Automatic geometry assessment against a master model	CAD models
2004	Shukur et al. [50]	Computer-based assessment	CAD drawings
2008	Covill et al. [16]	Online demonstrations	CAD models
2009	Fielke & Quinn [18]	E-portfolios	CAD models and drawing
2009	Paliokas [41]	Video tutorials and recordings	CAD models
2011	Menary et al. [36]	Interviews and video-audio submissions	CAD models and drawings
2012	Branoff et al. [7]	Rubrics	Feature-Based History-Based Parametric CAD models
2012	Sanna et al. [48]	Automatic geometry assessment against a master model	Polygonal models
2013	Devine & Laingen [17]	Rubrics	Design intent evaluation
2013	Irwin [25]	Scaffolding techniques	CAD models
2015	Company et al. [13]	Adaptable rubrics and assertion maps	Feature-Based History-Based Parametric CAD models
2016	Kirstukas [31]	Automatic assessment against a master model	Solid CAD Models
2017	Ramos-Barbero et al. [43]	Summaries of rules of design intent	Feature-Based History-Based Parametric CAD models

Tab. 1: Review of common CAD assessment approaches.

Baxter and Guerci developed a CAD assessment tool that compares geometric data from the student files to a master model provided by the instructor [4]. A similar approach was described by Kirstukas [31], where a computer program evaluates the geometry and alterability of student solid models based on sketched profiles, constraints, and mass properties against a model provided by the instructor [31]. However, both approaches emphasize the geometric aspects of the model and hardly assess design intent and strategic knowledge.

Alternatively, other authors have proposed the use of assessment as an active component of the training process. For example, Irwin [25] used scaffolding techniques (mentoring students toward finding solutions while adjusting the amount of support provided) to emphasize the importance of parameters to drive design intent and allow for increased flexibility of design exploration [25]. Likewise, Covill et al. [16] described a method of assessment where students were required to generate an online demonstration and describe their approaches and techniques to develop a CAD model.

Rubrics are common assessment tools in many disciplines used to articulate performance expectations for an assignment and evaluate student achievement [20],[21]. They can be defined as scoring instruments that list the criteria for evaluating a particular task along with gradations of achievement for each criterion [44]. Rubrics can be used to assess multi-dimensional performances [2] as well as complex performances [30],[38] and are often considered an invaluable tool for helping students understand the assessment criteria from the beginning [37].

The unique characteristics of rubrics make them suitable for performance evaluations as well as for supporting the process of formative assessment, where the rubrics regularly inform students about their progress and assist them throughout their learning process [6],[53]. Rubrics also aid in making students reflect on the quality of their own and other's work.

In the domain of CAD assessment, rubrics have been used successfully as an essential part of evaluation schemes to verify design intent [17]. Authors Company et al. [13],[14] discussed an approach to describe and evaluate CAD quality criteria in parametric models in the form of analytical rubrics (rubrics that break down a task into its components, which are evaluated separately and then combined to produce the overall score). Their solution was implemented as an expand-contract approach and used to introduce novice CAD users to high-level quality-oriented strategies early in their instruction. In these rubrics, the students' CAD model is compared against quality dimensions (conveyed as competences), and measured through evidences, which are expressed as "assertions [13]." Assertions maps are then 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.

This paper expands on previous research in the area of formative CAD rubrics for introducing CAD quality concepts at the early stages of instruction by providing a new assessment scheme for ~~top-down~~ bottom-up assembly modeling tasks. The term bottom-up is an assembly modelling strategy were individual parts are first modelled independently and then inserted into an assembly, where constraints are used to position the parts and describe interactions between parts. CAD quality dimensions for parametric models were redefined to accommodate the specific characteristics of the assembly modeling process. To validate this approach and further develop the proposed scheme, results of two studies are presented where inter-rater reliability between experts, and between experts and novices, were analyzed.

3 CAD ASSEMBLY RUBRICS

Taking the previous study on rubrics for parametric solid models [13] as a starting point, 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.

Descriptors can be understood as statements that communicate the desired state of each assessed aspect. Descriptors are defined by three primary characteristics: (1) they must correlate with a teachable result, (2) they must correlate with an unbiased and easily measurable result, and (3) 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 [33]. Achievement levels can be defined by two characteristics: (1) they must use the same terms as the corresponding criteria and (2) 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 [46], qualifiers can be described by frequency (never, rarely, sometimes, often, always, etc.), intensity (not at all, slightly, moderately, considerably, extremely, etc.), and probability (certainly not, unlikely, likely, certainly, etc.).

Figure 1 illustrates a sample portion of the assembly rubric as utilized in the experiments, showing the intensity qualifiers that describe the performance levels of a quality dimension.

CRITERION		WEIGHT	PERFORMANCE LEVELS				
#	**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 the linked files, require intense file management before they can be located and used	The file of the assembly, or the linked files, require moderate file management before they can be located and used	The file of the assembly, and the linked files, can be located and used but require minor file management	The file of the assembly, and the linked files, are easily located and used

Fig. 1: Sample portion of assembly rubric illustrating performance levels for criterion "the assembly is valid."

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. 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 multiple 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.)

A data quality dimension is defined as "a set of data quality attributes that represent a single aspect or construct of data quality [52]." This paper establishes the dimensions of the CAD quality space (as applied to assembly modeling) which are used to train novice CAD users similarly to the quality-oriented training approach described by Company et al [13] in the context of single part modeling. The proposed dimensions are inspired by the properties of representation schemes suggested by Requicha [45] and by the research team's previous work in the area of CAD model assessment. Formal properties include: domain, validity, completeness, and uniqueness. Informal properties are conciseness, ease of creation, and efficacy in the context of applications. A review of recent literature further supports this selection of properties. For example, model efficiency was emphasized by Bhavnani et al. [5] as well as Rynne and Gaughran [47] who also defined attributes for robustness and design intent. Similarly, Amadori et al. [1] discussed the importance of flexibility and robustness in CAD models.

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.

3.1 Validity

The original intent for this dimension 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 scores (so that maximum partial credit could be awarded). Of course, industry use of these rubrics may not benefit from such allowance. This Go/No/Go approach was recently validated by Company et al. [15] in an academic setting. The authors concluded that Go/No-Go criteria must be explicitly identified and included as such in the descriptor. 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 (including 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 does not contain 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 so that all files, including standard library parts, will be locally available.

3.2 Completeness

The completeness of a model refers to the inclusion of all the product aspects that are relevant for design purposes (i.e., they replicate the shape and size of the object) [13]. For single parts, completeness involves accurately defining the geometry and topology of the part by selecting the right set of geometric features and combining them properly. Similarly, an assembly can be considered complete if it includes all required components (parts and sub-assembly files), uses standard component files (library parts) when appropriate, and all components are correctly sized and 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 instances of an identical part, (see Figure 2).

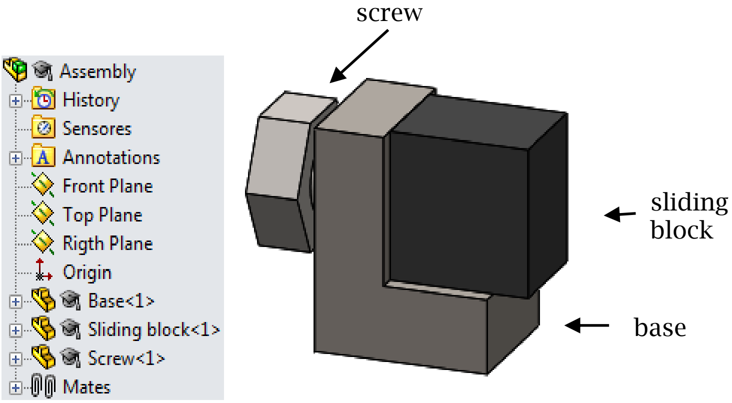


Fig. 2: Multiple instances of each part much match.

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 for rendering and presentation purposes (See Figure 3).

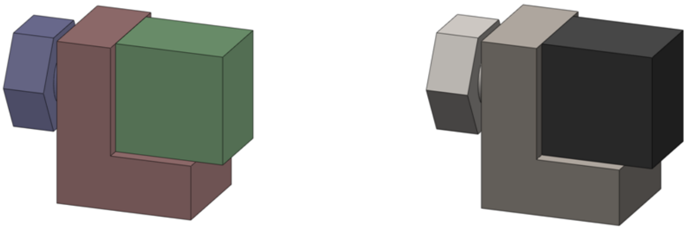


Fig. 3: Use of colors to differentiate components.

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 tools 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 (see Figure 4).

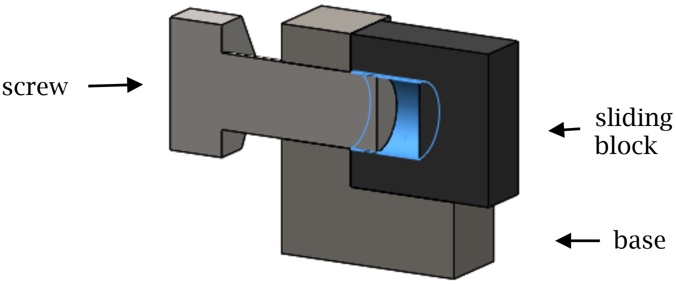


Fig. 4: Use of section to inspect correct placement.

Some software packages 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 symmetrically 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 (see Figure 5).

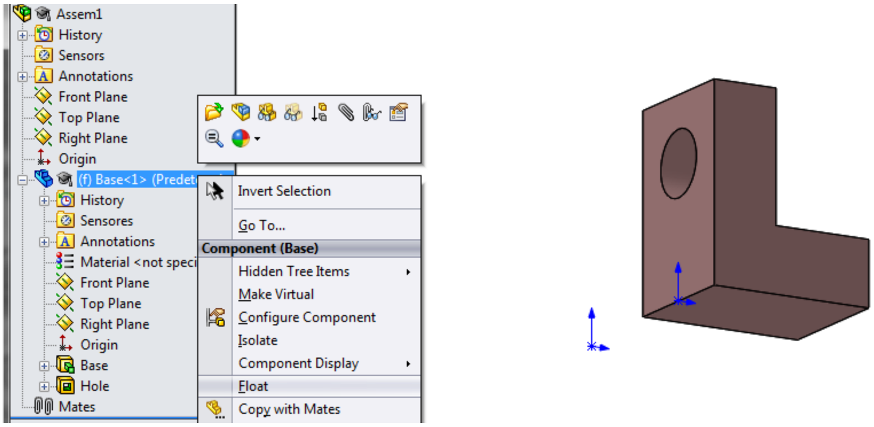


Fig. 5: Link base part to global reference system.

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.). As an example, it would be unnecessary to fully constrain a washer (Figure 6).

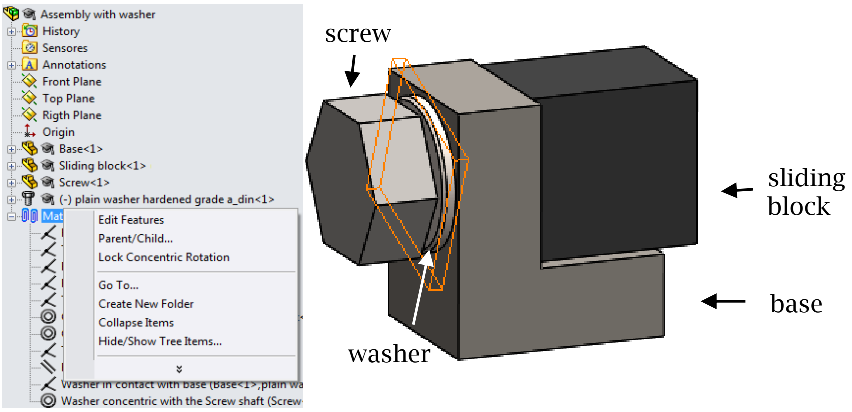


Fig. 6: Example of unnecessary motion constraint.

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 (see Figure 7).

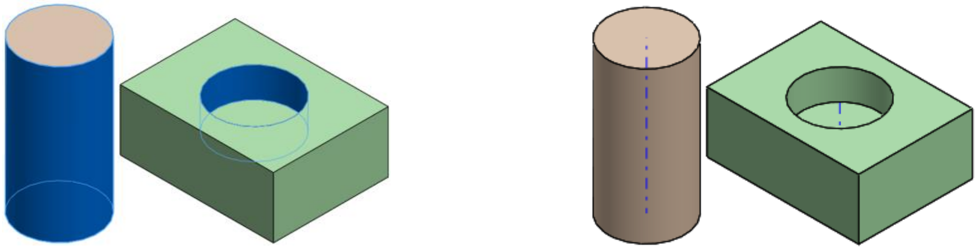
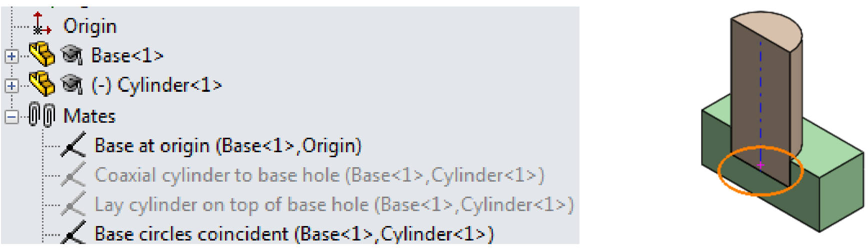


Fig. 7: If one cylinder is concentric with a hole (left), it is repetitive to mate both as coaxial (right).

Fragmented mates should also be avoided, as multiple simple mates are less efficient than one comprehensive one. As an example, the non-fragmented method to place a cylinder in a hole would be to mate the two contour circles (edges) as shown in Figure 8.



Origin
Base<1>
(-) Cylinder<1>
Mates
  Base at origin (Base<1>,Origin)
  Coaxial cylinder to base hole (Base<1>,Cylinder<1>)
  Lay cylinder on top of base hole (Base<1>,Cylinder<1>)
  Base circles coincident (Base<1>,Cylinder<1>)

Fig. 8: Non-fragmented mating procedure.

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.

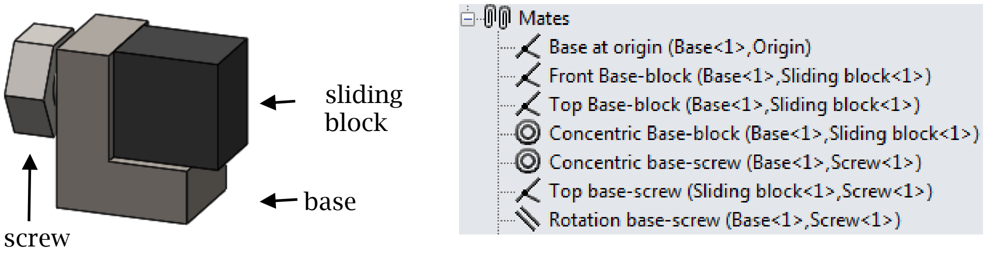
Highly semantic assembly operations provide context between user intentions and the contents of computational assemblies. These operations provide important design information that assist those that did not create the model to analyze and manipulate the assemblies. Good practices include using pattern 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 processes. For effective communication to occur, the document (assembly) must be easily understood (preferably at the first viewing) so other members of the design team can efficiently use, navigate, and understand the structure of the document and how it was built.

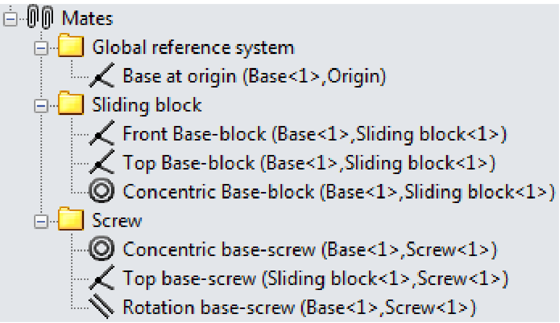
For single parts, clarity can be achieved by implementing good practices such as including labeling modeling operations in the model tree to emphasize their function, instead of how they were built; grouping related modeling operations in the model tree to emphasize parent-child relationships, following proper conventions [13]. In the context of assembly modeling, 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 recommended that all mates be re-labeled to emphasize just what exactly is linked, not the type of link implemented (see Figure 9).



- Mates
  - Base at origin (Base<1>,Origin)
  - Front Base-block (Base<1>,Sliding block<1>)
  - Top Base-block (Base<1>,Sliding block<1>)
  - Concentric Base-block (Base<1>,Sliding block<1>)
  - Concentric base-screw (Base<1>,Screw<1>)
  - Top base-screw (Sliding block<1>,Screw<1>)
  - Rotation base-screw (Base<1>,Screw<1>)

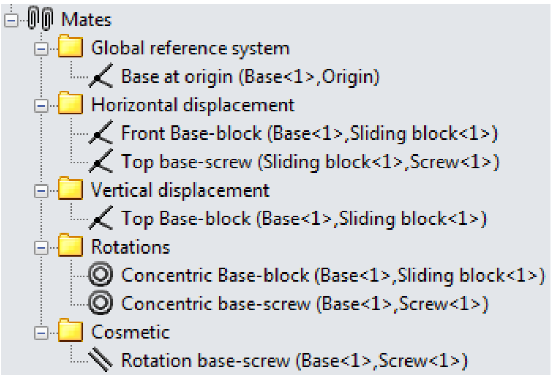
Fig. 9: Re-label mates to emphasize function.

Mating operations should be grouped according to the design criteria needed to increase communication. This process could be accomplished by grouping by parts (Figure 10) or degrees of freedom (Figure 11). While an optimum grouping procedure does not exist, it is more important to avoid clearly erroneous solutions.



- Mates
  - Global reference system
    - Base at origin (Base<1>,Origin)
  - Sliding block
    - Front Base-block (Base<1>,Sliding block<1>)
    - Top Base-block (Base<1>,Sliding block<1>)
    - Concentric Base-block (Base<1>,Sliding block<1>)
  - Screw
    - Concentric base-screw (Base<1>,Screw<1>)
    - Top base-screw (Sliding block<1>,Screw<1>)
    - Rotation base-screw (Base<1>,Screw<1>)

Fig. 10: Mating operations grouped by parts.



- Mates
  - Global reference system
    - Base at origin (Base<1>,Origin)
  - Horizontal displacement
    - Front Base-block (Base<1>,Sliding block<1>)
    - Top base-screw (Sliding block<1>,Screw<1>)
  - Vertical displacement
    - Top Base-block (Base<1>,Sliding block<1>)
  - Rotations
    - Concentric Base-block (Base<1>,Sliding block<1>)
    - Concentric base-screw (Base<1>,Screw<1>)
  - Cosmetic
    - Rotation base-screw (Base<1>,Screw<1>)

Fig. 11: Mating operations grouped by degrees of freedom.

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 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. In the context of parametric solid modeling of single parts, design intent addresses the proper planning of the model structure and the effects of changes in the model by altering one or more dimensional constraints. In assembly modeling, design intent refers to the proper planning and sequence of the assembly constraints and the behavior of the assembly structure when parts are modified, and the entire assembly model must be rebuilt. In this regard, the idea of design intent at the assembly level is closely related to design intent at the individual part level, as altering one part may have unintended consequences in the overall structure of the assembly (e.g., invalid assembly constraints, missing references, incorrect placements, etc.). In our paper, however, the effects of these interconnections between design intents of parts and the assemblies as well as unintentional effects that may occur when assembly components are modified in-context are not included in the proposed rubric.

Many design methodologies use assembly models to investigate design behavior. Assembly models convey design intent when they convey information that is useful for analyses. Four different aspects can be analyzed: (1) Assembly planning, when assembly sequence is paramount; (2) Assembly Process Design (APD), when functionalities are examined; (3) Design for Assembly (DFA), when affordances used to assemble and disassemble are analyzed; and (4) 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 (see Figure 12).

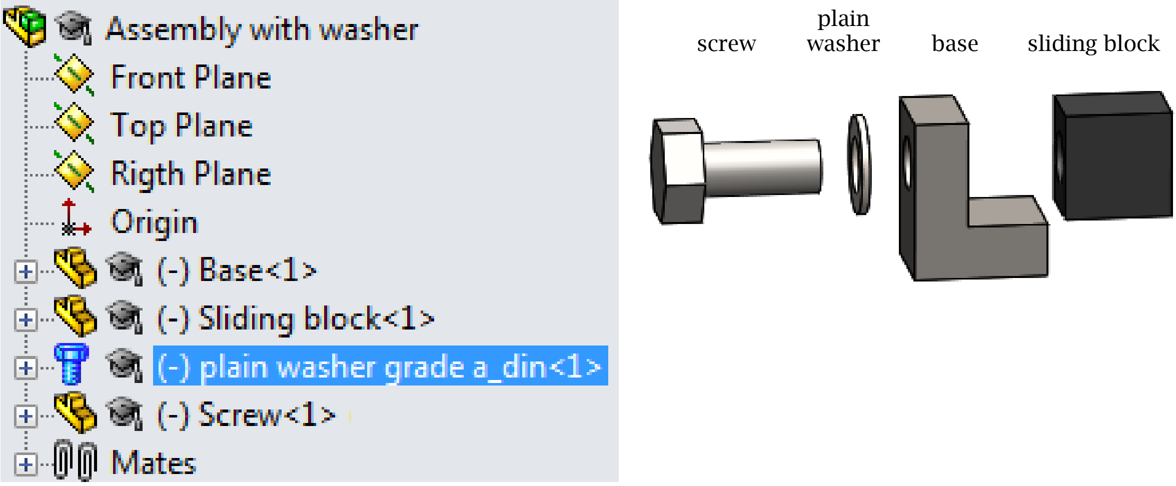
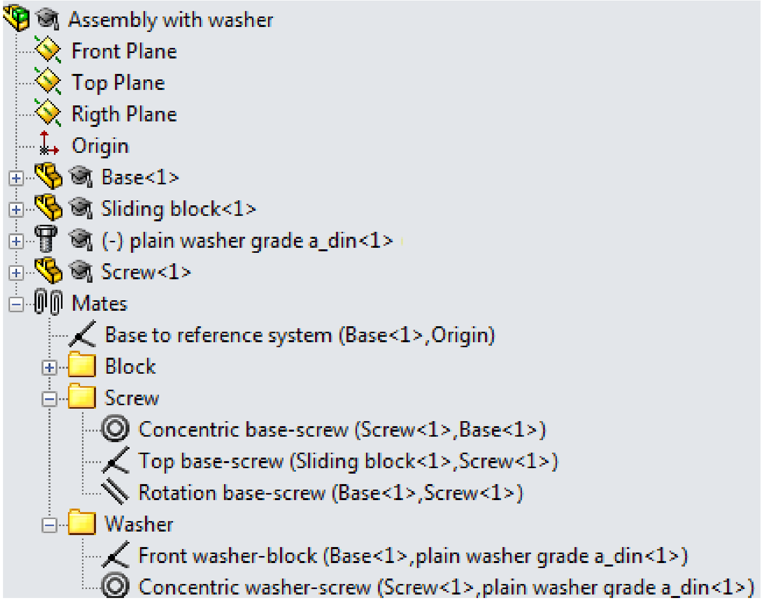


Fig. 12: Assembly sequencing from main to auxiliary elements.

A realistic assembly sequence could result in unrealistic mates, so agreement between actual sequencing and reasonable mate linking is imperative (see Figure 13).



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>)

Fig. 13: Example of realistic mating sequence causing unrealistic mating situation. Notice that the washer is inserted before the screw, but is linked to the screw, which was inserted afterward.

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. DFA guidelines and best practices are particularly relevant to how the dimension of design intent can be implemented at the assembly level, as relationships between physical components can often be described by analogous constraints in the assembly model. In this regard, 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 independence of replaceable parts should be increased.

To summarize, the assertions for assembly modeling are shown in Table 2.

Code	Description
1	**The assembly is valid (soft Go/No-Go criterion, which multiplies 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.
1.1a	The file of the assembly has the expected contents (and name) and is in the expected place (folder or website).
1.1b	The file of the assembly can be re-opened after closing the current session (even on a different computer).
1.1c	The file of the assembly opens in a neutral state (i.e. no operations in progress).
1.2	The assembly can be used.
1.2a	The assembly is compatible with the CAD program (and software version) used by the receiver.
1.2b	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 compatibility issues exist between versions.
1.3a	All parts linked to the assembly can be accessed.
1.3b	All sub-assemblies linked to the assembly can be accessed.
1.3c	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).
2.1a	The assembly includes all the components and their corresponding copies.
2.1b	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.
2.2a	Standard library components are used when required.
2.2b	Standard library components are suitably instantiated from the library.
2.3	Components (parts, sub-assemblies, and library components) are correctly placed.
2.3a	Relative locations among components match their functional positions.
2.3b	Components are free of unwanted interferences.
3	**The assembly is consistent.**
3.1	The base component is correctly assigned and is well linked to the global reference system.
3.1a	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).
3.1b	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.
3.2a	Assembly mates prevent invalid movement.
3.2b	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.
4.2a	3D patterns operations (translate-and-repeat, rotate-and-repeat) are used whenever possible.
4.2b	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.
5.1a	Components are labeled and grouped to emphasize their function, instead of how they were defined.
5.1b	Mates are labeled to emphasize their function.
5.1c	Related mates are grouped to emphasize parent/child relationships.
5.2	The assembly uses compatible and standard mates.
5.2a	The most compatible mates are always used.
5.2b	The most standard mates are always used.
6	**The assembly conveys design intent**
6.1	The assembly tree replicates the assembly/disassembly process
6.1a	The assembly sequence proceeds from main to auxiliary elements.
6.1b	The assembly sequence reflects a realistic mounting sequence.
6.2	Sub-assemblies have been properly identified and efficiently used
6.2a	Sub-assemblies encapsulate clearly perceived functions.
6.2b	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.
6.3a	Mating features provided to grasp, move, orient, and insert the part, if any, have been identified.
6.3b	Mating features provided to grasp, move, orient, and insert the part, if any, are mostly used for mating.

Tab 2: Assertions for assembly modeling.

3.7 Rating Scale

To accommodate the 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 dimensions); Complete: 20%; Consistent: 30%; Concise: 20%; Clear: 15%; and Design Intent: 15%.

As stated previously, Validity was designated as Go/No-Go criteria, with this dimension (along with the concurrent sub-dimensions) engendering a total weight percentage of 0%. Completeness, worth 20% of the total score, was reflected by each of the three sub-dimensions equaling a weight of 6.67%. The two sub-dimensions of Consistency (worth 30%), was allotted 15% each. The three sub-dimensions of Conciseness (30% total) were valued at 6.67% each. Clarity (15% total) was defined by two sub-dimensions rated 7.5%. Design Intent (15% total weight), was comprised by three sub-dimensions equally divided. These sub-dimensions, and their explanations, were discussed previously in Table 2.

Rohrmann [46] 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; and 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.

4 EXPERIMENTAL STUDY

Two experiments were conducted assessing student understanding of assembly rubrics using our custom software platform "Annota e-Rubrics" [12]. 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.

4.1 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 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, these quality dimensions were aligned with preceding research on parts rubrics, accomplished by our research team [13].

Completion of Annota rubrics was 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. The evaluators came from institutions with different teaching styles and curricula.

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 custom 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 are an L-Bracket, Bolt, Base, and Wheel (see Figure 14). The objects modeled for this experiment and the specific curricular concepts addressed were selected by the instructor of record based on course requirements and the professional judgment of the instructor.

Two sub-assemblies were assumed: (a) Anchor Arm consisting of Base, L-Brackets, and Fixing Screws and (b) Bearing Wheel consisting of Wheel and Bearing. The expected sub-assemblies are shown in Figure 15, with the final assembly solution provided in Figure 16. 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).

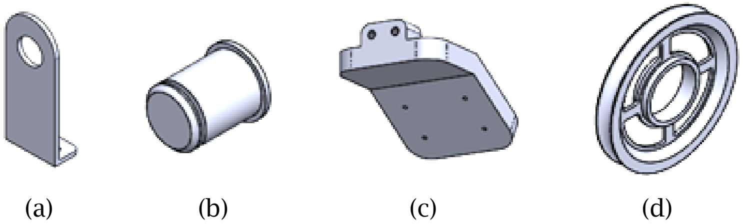

(a) (b) (c) (d)

Fig. 14: Non-standard parts used for modeling in Experiment 1: (a) L—Bracket, (b) Bolt, (c) Base, and (d) Wheel.

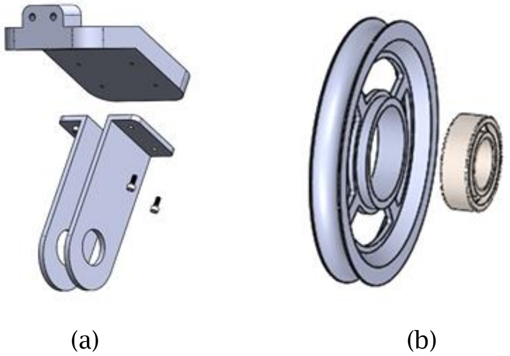

(a) (b)

Fig. 15: Sub-assemblies used in Experiment 1: (a) Anchor arm and (b) Bearing wheel.

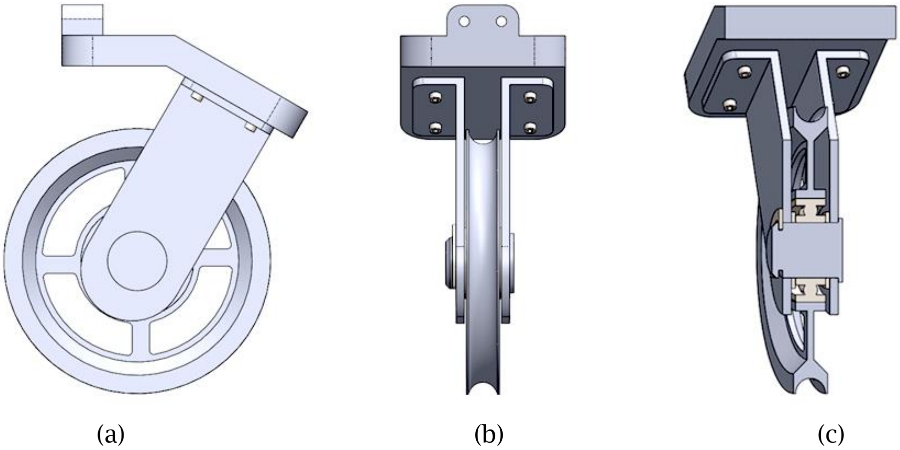

(a) (b) (c)

Fig. 16: Front (a), side (b), and sectional (c) views of the final assembly used in Experiment 1.

4.2 Experiment 1 Discussion

Table 3 shows, for each quality criteria, the difference in scores between instructors, and the difference in scores between the instructor average and each student.

Assembly Dimensions	**Between Instructors**	**Instructors vs. Students**
Dim 1: Validity	-0.1	-0.15
Dim 2: Completeness	-0.02	0.12
Dim 3: Consistency	-0.03	0.03
Dim 4: Conciseness	-0.03	-0.05
Dim 5: Clarity	-0.04	0.03
Dim 6: Design Intent	-0.11	-0.28

Tab. 3: Experiment 1: Differences in scores between instructors and between the instructor average and students (Max score = 5).

For the dimension of Validity, instructors were more pessimistic than students, but both instructors provided similar assessments. For the dimension of Completeness, 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%), Conciseness (within 5%), and Clarity (within 4%). Both instructors were strongly more pessimistic than the students for the dimension of Design Intent, which is to be expected as it is a difficult concept for beginning learners to grasp.

The first hypothesis was that assembly rubrics produce an objective accumulative assessment of students. In order to validate this hypothesis, the assessment performed by Instructor 1 was compared against the assessment made by Instructor 2 (instructor inter-rater reliability 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 [23]. The advantage is that if inter-rater reliability is high, raters can be used interchangeably [23], 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 4 illustrates the inter-rater reliability scores for Experiment 1 (for the student and both instructors). At first glance, it can be 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.

% agreement	**Dim. 1 (Valid)**	**Dim. 2 (Complete)**	**Dim. 3 (Consistent)**	**Dim. 4 (Concise)**	**Dim. 5 (Clear)**	**Dim. 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

Tab. 4: Interrater reliability scores in Experiment 1.

Table 5 displays the Pearson Correlation values for Experiment 1 (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.

Correlation Coefficient	**Dim. 1 (Valid)**	**Dim. 2 (Complete)**	**Dim. 3 (Consistent)**	**Dim. 4 (Concise)**	**Dim. 5 (Clear)**	**Dim. 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

Tab. 5: Pearson Correlation values in Experiment 1.

Results from this experimental study 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 [30]. 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 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 6), other than Dimension 6, where the medians were significantly different. For clarity, values less than 0.05 (greater differences in medians) are shaded, 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.

Observation relationship	**Dim. 1 (Valid)**	**Dim. 2 (Complete)**	**Dim. 3 (Consistent)**	**Dim. 4 (Concise)**	**Dim. 5 (Clear)**	**Dim. 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

Tab. 6: P-Values for Wilcoxon Signed Rank Test in Experiment 1.

4.3 Experiment 2

Following a similar procedure as the first experiment, Experiment 2 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 custom 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 is shown in Figure 17.

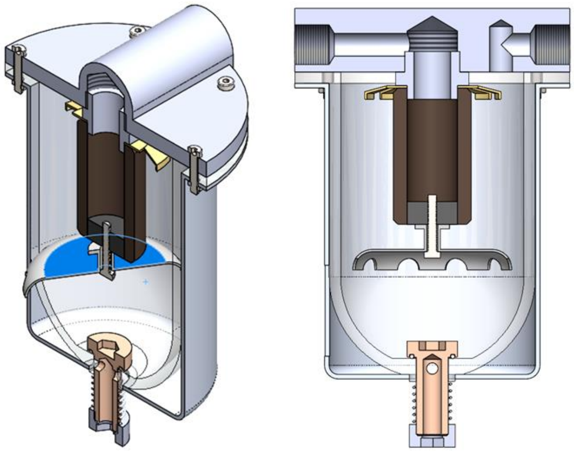


Fig. 17: Sectioned filter assembly used in Experiment 2. Isometric (left) and Front View (right).

4.4 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.

Table 7 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 students.

Assembly Dimensions	between instructors	instructors vs. students
Dim 1: Validity	0.0	-0.02
Dim 2: Completeness	-0.02	0.04
Dim 3: Consistency	-0.07	-0.07
Dim 4: Conciseness	-0.02	0.05
Dim 5: Clarity	-0.02	0.10
Dim 6: Design Intent	-0.09	-0.10

Tab. 7: Experiment 2: Differences in scores between instructors and between the instructor average and students (Max score = 5).

For the dimension of validity, 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 4%), consistency (within 7%), and conciseness (within 5%). For the dimension of clarity, instructors were more optimistic than the students, and the instructor assessments were close. For the dimension of design intent, the instructors were more pessimistic than the students.

Table 8 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 measurable increase in agreement for all dimensions other than validity, for instructors and students, between Experiment 1 and Experiment 2. 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.

% agreement	Dim. 1 (valid)	Dim. 2 (complete)	Dim. 3 (consistent)	Dim. 4 (concise)	Dim. 5 (clear)	Dim. 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

Tab. 8: Interrater reliability scores in Experiment 2.

Table 9 shows the Pearson Correlation values for Experiment 2 (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 Experiment 1, 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.

Correlation Coefficient	**Dim. 1 (valid)**	**Dim. 2 (complete)**	**Dim. 3 (consistent)**	**Dim. 4 (concise)**	**Dim. 5 (clear)**	**Dim. 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	0.57510	0.60329	0.48355
Instructor 1-Instructor 2	1	0.98461	0.97757	0.96389	0.98004	0.95149

Tab. 9: Pearson Correlation values in Experiment 2.

Table 10 shows the p-values for the Wilcoxon Signed Rank Test. Once again, for clarity, values less than 0.05 (greater differences in medians) are shaded and values greater than 0.05 (more similar medians) remain unshaded. Values that could not be computed are also indicated. 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.

Observation relationship	**Dim. 1 (valid)**	**Dim. 2 (complete)**	**Dim. 3 (consistent)**	**Dim. 4 (concise)**	**Dim. 5 (clear)**	**Dim. 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

Tab. 10: P-Values for Wilcoxon Signed Rank Test in Experiment 2.

5 DISCUSSION

Individual results from our two experimental studies were compared side by side to determine overall correlations. Table 11 shows whether each dimension increased, decreased, or remained unchanged in respect to rubric understanding, between Experiment 1 and Experiment 2 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.

Difference	**Dim. 1 (valid)**	**Dim. 2 (complete)**	**Dim. 3 (consistent)**	**Dim. 4 (concise)**	**Dim. 5 (clear)**	**Dim. 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

Tab. 11: Percent Agreement values between Experiment 1 and Experiment 2.

Table 12 reflects whether each dimension increased, decreased, or remained unchanged in respect to rubric understanding, between Experiment 1 and Experiment 2 for the Pearson Correlation values. Dimensions 1, 2, 4, and 5 showed increased correlation, but decreased correlation is reported for Dimensions 3 and 6.

Difference	**Dim. 1 (valid)**	**Dim. 2 (complete)**	**Dim. 3 (consistent)**	**Dim. 4 (concise)**	**Dim. 5 (clear)**	**Dim. 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**

Tab. 12: Pearson Correlation values between Experiment 1 and Experiment 2.

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 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 students 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 FUTURE WORK

Appraisal of student performance is a critical component necessary for engaged student learning. The use of rubrics to perform CAD 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.

The results of these experiments with assembly rubrics reveal 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, but little correlation exists for the dimensions of consistency and design intent. Probable reasons for the lack of correlation for these two dimensions could be attributed to the fact that they involve more complex modeling concepts, consisting of intimate knowledge of: (1) position of the model in reference to various reference systems, (2) understanding of proper and improper movement of components within the assembly, and (3) purpose of the mechanism, how it functions, and which components are needed as anchors within the overall model.

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.

In general, the modest differences between instructors suggest that the proposed strategy 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, further refinement is necessary to formally determine the validity to provide formative self-evaluation of CAD assembly skills for new learners. 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. The research team surmises that improving the instructional materials (primarily in the dimensions of consistency and design intent) and increasing rubric exposure are perhaps, the first steps to obtain valid formative self-evaluation of CAD assembly skills for beginning users.

More easily understood introductory material for students could possibly increase comprehension of various misunderstood rubric dimensions, but these concepts may be 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. 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.

Finally, assembly rubrics (especially if offered in electronic formats) offer the possibility of them being integrated into CAD software tools, facilitating the automation of the assessment process. These tools could be used to not only guide the CAD operator in the design process, but could also assess the model undergoing creation. 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.

7 ACKNOWLEDGEMENTS

This work was partially supported by the Spanish grant DPI2017-84526-R (MINECO/AEI/FEDER, UE), project "CAL-MBE, Implementation and validation of a theoretical CAD quality model in a Model-Based Enterprise (MBE) context.", and the ANNOTA2 project funded by Universitat Politècnica de València.

Jeffrey Otey, http://orcid.org/[0000-0002-3763-8759]
Pedro Company, http://orcid.org/[0000-0001-6399-4717]
Manuel Contero, http://orcid.org/[0000-0002-6081-9988]
Jorge D. Camba, http://orcid.org/[0000-0001-5384-3253]

REFERENCES

[1] Amadori K, Tarkian M, Ölvander J, Krus P. Flexible and robust CAD models for design automation. Adv Eng Inf 2012;26(2):180–95. https://doi.org/10.1016/j.aei.2012.01.004

[2] Andrade, H.; Valtcheva, A.: Promoting learning and achievement through self-assessment. Theory into Practice 2009;48(1):12–19. https://doi.org/10.1080/00405840802577544

[3] ASME Y14.41-2012 Digital Product Definition Data Practices. The American Society of Mechanical Engineers, 2012, New York.

[4] Baxter, DH; Guerci, MJ.: Automating an introductory computer aided design course to improve student evaluation. Presented at the 2003 ASEE Annual Conference & Exposition, Nashville, Tennessee. https://peer.asee.org/11479

[5] Bhavnani SK.; John BE.; Flemming U.: 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. Pittsburgh (USA, New York): ACM Press; 1999:183–190. https://doi.org/10.1145/302979.303036

[6] Black, P.; Wiliam, D.: Developing the theory of formative assessment. Educational Assessment Evaluation and Accountability 2009;21:5–31. https://doi.org/10.1007/s11092-008-9068-5

[7] Branoff, TJ.; Dobelis, M.: Engineering graphics literacy: Measuring students' ability to model objects from assembly drawing information. In Proceedings of the 66th Midyear Conference of the Engineering Design Graphics Division of the American Society for Engineering Education 2012;22-24.

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

[9] Brunnermeier SB.; Martin SA.: Interoperability costs in the US automotive supply chain. Supply Chain Manag: Int J 2002;7(2):71–82. https://doi.org/10.1108/13598540210425821

[10] Chester, I.: Teaching for CAD expertise. International Journal of Technology and Design Education 2007;17(1):23-35. https://doi.org/10.1007/s10798-006-9015-z

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

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

[13] Company P.; Contero M.; Otey J.; Plumed R.: Approach for developing coordinated rubrics to convey quality criteria in MCAD training. Computer-Aided Design 2015;63:101–117. https://doi.org/10.1016/j.cad.2014.10.001

[14] Company P.; Contero M.; Salvador-Herranz G.: Testing rubrics for assessment of quality in CAD modelling. In: REES 2013. Proceedings of the research in engineering education symposium 2013;107–112.

[15] Company, P.; Otey, J.; Agost, MJ.; Contero, M.; Camba, JD. Go/No Go Criteria in Formative E-Rubrics. In International Conference on Learning and Collaboration Technologies (pp. 254-264). 2018, Springer, Cham.

[16] Covill, D.; Katz, T.; Morris, R.: Teaching and assessing CAD using online demonstrations. In DS 46: Proceedings of E&PDE 2008, the 10th International Conference on Engineering and Product Design Education, Barcelona, Spain, 04.-05.09. 2008.

[17] Devine, KL.; Laingen, MA.: Assessing design intent in an introductory-level engineering graphics course. Presented at the 68th Mid-Year Conference of the ASEE Engineering Design Graphics Division 2013;59-63.

[18] Fielke, JM.; Quinn, D.: Supporting self-assessment using e-portfolios. In 20th Annual Conference for the Australasian Association for Engineering Education, 6-9 December 2009: Engineering the Curriculum (p. 405). Engineers Australia.

[19] González-Lluch, C.; Company, P.; Contero, M.; Camba, JD.; Plumed, R.: A survey on 3D CAD model quality assurance and testing tools. Computer-Aided Design 2017;83:64-79. https://doi.org/10.1016/j.cad.2016.10.003

[20] Goodrich Andrade H.: Understanding rubrics. Educ Leadersh 1996;54(4):14–7.

[21] Goodrich Andrade H.: Using rubrics to promote thinking and learning. Educ Leadersh 2000;57(5):13–8.

[22] Gu H.; Chase TR.; Cheney DC.; Bailey T.; Johnson D.: Identifying, correcting, and avoiding errors in computer-aided design models which affect interoperability. Journal of Computing and Information Science in Engineering 2001;1(2):156–166. https://doi.org/10.1115/1.1384887

[23] Gwet, KL.: Handbook of inter-rater reliability (4th ed.). Gaithersburg, MD: Advanced Analytics, LLC, 2014.

[24] Hartman, NW.: The development of expertise in the use of constraint-based CAD tools: Examining practicing professionals. Engineering Design Graphics Journal 2009;68(2).

[25] Irwin, J.: Wind turbine blade CAD models used as scaffolding technique to teach design engineers. Engineering Design Graphics Journal 2013;77(2):1-17.

[26] ISO 10303-59. Industrial automation systems and integration -- Product data representation and exchange- Part 59: Integrated generic resource: Quality of product shape data. Geneva (Switzerland): International Organization for Standardization, 2014.

[27] ISO 16792:2015 Technical Product Documentation – Digital Product Definition Data Practices. Organisation Internationale de Normalisation, Genève, Suisse, 2015.

[28] Jackson C.; Buxton M.: The design reuse benchmark report: seizing the opportunity to shorten product development. Benchmark report. Boston (MA): Abeerden Group (US); 2007. Available from URL: http://www.aberdeen.com/summary/report/benchmark/DesRes_Out_3908.asp

[29] Johnson, RA.; Miller, I.; Freund, JE.: Probability and statistics for engineers. Pearson Education, 2000.

[30] Jonsson, A.; Svingby, G.: The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review 2007;2:130–144. https://doi.org/10.1016/j.edurev.2007.05.002

[31] Kirstukas, SJ.: Development and evaluation of a computer program to assess student CAD models. Proceedings of ASEE's 123rd Annual Conference and Exposition, New Orleans, LA, 2016. https://doi.org/10.18260/p.26781

[32] Leahy, K.: 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, Limerick, Ireland 2012;59-66.

[33] Likert, R.: A technique for the measurement of attitudes. Archives of psychology, 1932.

[34] Lubell J.; Chen K.; Horst J.; Frechette S.; Huang P.: Model Based Enterprise. Technical Data Package Summit Report. NIST Technical Note 1753, 2012.

[35] McKenney D.: Model quality: the key to CAD/CAM/CAE interoperability. In: Proceedings of the 1998 MSC software Americas users conference. Universal City, CA.

[36] Menary, G.; Robinson, T.; Belfast, NI.: Novel approaches for teaching and assessing CAD. In International Conference on Engineering Education, Belfast, N. Ireland, 2011;21-26.

[37] Montgomery K.: Authentic tasks and rubrics: going beyond traditional assessments in college teaching. Coll Teach 2002;50(1):34–40. https://doi.org/10.1080/87567550209595870

[38] Moskal, BM.; Leydens, JA.: Scoring rubric development: Validity and reliability. Practical Assessment, Research & Evaluation 2000;7(10).

[39] Otey, J.; Company, P.; Contero, M.; Camba, JD.: A Review of the Design Intent Concept in the Context of CAD Model Quality Metrics. In: ASEE Annual Conference and Exposition in Indianapolis, IN. 2014.

[40] Otey, J.; Company, P.; Contero, M.; Camba, JD.: Revisiting the design intent concept in the context of mechanical CAD education. Computer-Aided Design and Applications 2017;15(1):1-14. https://doi.org/10.1080/16864360.2017.1353733

[41] Paliokas, I.: Reinforcing metacognition in CAD education using videotutorials. Computer-Aided Design and Applications 2009;6(5):613-623. https://doi.org/10.3722/cadaps.2009.613-623

[42] Ramos-Barbero, B.; Garcia-Garcia, R.: 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 2009;19(1):146-160. https://doi.org/10.1002/cae.20300

[43] Ramos-Barbero, B.; Melgosa-Pedrosa, C.; Zamora-Samperio, R.: Learning CAD at university through summaries of the rules of design intent. International Journal of Technology and Design Education 2017;27(3):481-498. https://doi.org/10.1007/s10798-016-9358-z

[44] Reddy, YM.; Andrade, H.: A review of rubric use in higher education. Assessment & Evaluation in Higher Education 2010;35(4):435–448. https://doi.org/10.1080/02602930902862859

[45] Requicha AAG.: Representations for rigid solids: theory, methods, and systems. ACM Comput Surv 1980;12(4):437–64. https://doi.org/10.1145/356827.356833

[46] Rohrmann, B.: Verbal qualifiers for rating scales: Sociolinguistic considerations and psychometric data. Project Report. University of Melbourne, Australia, 2007.

[47] Rynne, A.; Gaughran, W.: Cognitive modeling strategies for optimum design intent in parametric modeling. Computers in Education Journal 2008;18(1):55-68.

[48] Sanna, A.; Lamberti, F.; Paravati, G.; Demartini, C.: Automatic assessment of 3D modeling exams. IEEE Transactions on Learning Technologies 2012;5(1):2-10. https://doi.org/10.1109/TLT.2011.4

[49] SASIG, 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)

[50] Shukur, Z.; Away, Y.; Dawari, MA.: Computer–Aided Marking System for Engineering Drawing. In Society for Information Technology & Teacher Education International Conference. Association for the Advancement of Computing in Education (AACE), 2004;1852-1857.

[51] Son S.; Na S.; Kim K.: Product data quality validation system for product development processes in high-tech industry. International Journal of Production Research 2011;49(12):3751–3766. https://doi.org/10.1080/00207543.2010.486906

[52] Wang, RY.; Strong, DM.: Beyond accuracy: what data quality means to data consumers. J Manage Inf Syst 1996;12(4):5–34. https://doi.org/10.1080/07421222.1996.11518099

[53] Wiliam, D.: What is assessment for learning? Studies in Educational Evaluation 2011;37:2–14. https://doi.org/10.1016/j.stueduc.2011.03.001

[54] Yang, J.; Han, S.; Kang, H.; Kim, J.: Product data quality assurance for e-manufacturing in the automotive industry. International Journal of Computer Integrated Manufacturing 2006;19(2):136–147. https://doi.org/10.1080/09511920500171261

[55] Yang, J.; Han, S.: Repairing CAD model errors based on the design history. Computer Aided Design 2006;38(6):627–640. https://doi.org/10.1016/j.cad.2006.02.007