TOWARDS COMPUTATIONAL TOOLS FOR SUPPORTING THE REFLECTIVE TEAM

ANDREW W. HILL, ANDY DONG, AND ALICE M. AGOGINO
University of California at Berkeley
5136 Etcheverry Hall
Berkeley, CA 94720-1750
{ahill,adong,aagogino}@me.berkeley.edu

Abstract. The content of engineering design documentation, beyond capturing the details of the design, communicates the shared knowledge integration of the design team. In this research, we present a method to analyze design documentation for levels of *shared understanding* and *team cohesiveness* in design teams by applying the computational tools of latent semantic analysis and natural language processing. We study the design documentation from students in a multidisciplinary, graduate-level product design and development course. The results show promise in measuring cohesiveness and shared understanding in design teams by analyzing their documentation and correlating metrics of effective communication to the likelihood of successful outcomes of the design team.

1. Introduction

1.1 MOTIVATION

A dialectic exists between two prevailing paradigms of design: Simon's (Simon, 1996) rational problem solving process and Schön's (Schön, 1983) reflective practice. Whereas Simon's viewpoint is prescriptive and positivistic, Schön's is constructivist and phenomenological (Dorst, 1997). In solving a complex design problem, the rational problem solving practitioner might propose the decomposition of the design problem into reasonably independent subproblems for which a solution could be built from known scientific and engineering principles (Pahl and Bietz, 1996), whereas the reflective practitioner might suggest re-framing the problem in a way that opens up the possibility of moving in a new design direction.

While, in the context of this paper, it is not necessary to ascribe to a particular theory of design, we cite these theories to highlight the roles of rational problem solving and reflection in design. Although reflective practice may better describe design teams in practice (Valkenburg, 2000), in collaborative, team-based design environments, successful design outcomes require both rational problem solving and the knowledge integration (shared understanding) of the design team. High-performance teams excel in developing shared understanding (Cooper and Kleinschmidt, 1995; Griffin, 1995) through cooperative exchange of information and mutual agreements (Citera et al., 1995). Social interaction in the design process is a significant determinant of the success of collaborative design (Bucciarelli, 1994) and is likely as equal a determinant as the expertise of the team.

Cross-functional teams exist in nearly all firms engaged in design, such as product development and architecture (Henke et al., 1993). Ideally, at each stage of the design process, a team would be able to draw upon and reflect upon all of their collective knowledge to determine the best course of action. But the very *raison d'être* of cross-functional design teams – bringing the required skills to solve design problems "upstream" in the design process – is also the primary challenge of design teams. Incompatible viewpoints among design team members and failure to negotiate different design perspectives and specialties may result in ineffective collaboration. In practice, numerous social barriers exist, such as ineffective leadership, lack of shared objectives and cultural heterogeneity. Without a strong product champion or manager to detect transgressions and facilitate their resolution, disagreements could lead to a break down of the team. Helping teams to form a shared understanding is an important management role.

However, with the increasing trend to virtual and distributed teams, the manager and all team members may not be available for in-person monitoring and team self-evaluation. While numerous computer-supported collaborative work tools exist to foster collaboration and enhance rational problem solving, few tools exist to help teams reflect on their process. Part of the reason for the lack of these tools is that formal methods for cognitive modeling of the psychosocial behavior of design teams have not been extensively examined. Advancements in understanding how design teams acquire and maintain their collective identity would improve our ability to build tools that assist teams to reflect on their process.

For this reason, we have been studying computational methods to better understand the attitudes and behavior of teams in real-time in order to build tools that will have an evaluative impact on dealing with the nuances of the interactions as they occur and provide mechanisms for constructive in-process improvements of team performance. We are developing reflective, evaluative tools for individuals seeking to improve effective communication in their teams so that the communication may lead to increased shared understanding and cohesiveness.

In this paper, we studied the documentation of a group of nine collaborative design teams, collectively totaling 37 men and 7 women, in the context of a multidisciplinary, graduate course in product design and development. The student teams attend a course that emphasizes customer-driven product development and blends the study of design theories and methods with execution of a product development project[1]. Students from computer science, engineering, business and information management and systems join forces on product development teams of four to five members coached by faculty and professional designers from industry. Because the students self-select both the product and the make-up of the team, each student brings his or her own disciplinary perspective to the team effort, and must learn to synthesize that perspective with those of the other students in the group to develop a sound, marketable product. Part of the learning in this course is for the students to assess patterns of cooperation and team dynamics and to reflect on both the behavioral and organizational challenges the teams faced. Large amounts of customer feedback and user input are encouraged through the process and generally teams do a quality job of this.

Within each team, the students' worked collaboratively on their design product and project deliverables, which consisted primarily of design documentation and reflection on the team's execution of the process at each stage of the product development process. We used these documents to measure the shared understanding and cohesiveness of the design teams. The teams utilized the course management system from Blackboard, Inc. to share design documents with each other and the faculty. Design documents were captured and analyzed using the computational linguistic tools of latent semantic analysis (LSA) (Landauer et al., 1998) and natural language processing. We developed and tested a method to ascertain the level of shared understanding and cohesion of these teams based on their design documentation.

1.2 THEORETICAL FRAMEWORK AND HYPOTHESES

To a large extent, the theoretical underpinnings for our research in teams rest on social network theory and social interaction theory (Adams, 1967; Cook and Whitmeyer, 1992; Wasserman and Galaskiewicz, 1994). Questions surrounding social cohesion (Moody and White, 2000) – understanding the foundations of group formation and interaction – have engaged sociologists extensively. Relationships within these social networks have as an integral element "intimate communication." The purpose of

[1] http://best.me.berkeley.edu/~aagogino/me290p/me290p.html

communication in a team is to establish a set of coherent ideas and "get them across" to design team members and stakeholders. As opposed to a merely passive purpose, namely the passive transmission and reception of information, the communication serves an active role of generating new meaning. Communication in a social setting is often characterized as the creation of shared understanding through interaction among people.

Groups of people who communicate often and over long periods of time form "semantic communities" (Robinson and Bannon, 1991) with their own conventions of meaning. Similarly, Bødker and Pedersen (1991) define the term "workplace cultures" to describe a group of workers that share a common "code of conduct." Research in information use in design (Baird et al., 2000; Lloyd, 2000) show that words and phrases used by designers in the design process often captured personal experience and contributed to a wider narrative at the team, project or corporation level.

Because design knowledge is generated, codified, and reflected on in an ongoing process between the various stakeholders, teams must be able to synthesize this shared knowledge into a shared understanding in order to assure successful outcomes. Our approach to identifying the correlation between design documents and shared understanding is to detect the underlying patterns of word choices and meaning that express the shared understanding of a group of communicators. The premise is that textual coherence is both an indicator and measure of team cohesiveness when that text is a product of intra-team communication.

The ability of a team to work cohesively together towards a common goal plays a large role in the overall success of the team. There are many factors that affect the *cohesiveness* of a team such as commitment to the task, group pride and interpersonal relationships. Nonetheless, there is a significant correlation between cohesiveness and performance (Neck and Manz, 1994). Observational studies and surveys of teams have shown that many of the problems that occur in a team arise from team members (or the entire team itself) becoming lost or diverging (McDonough et al., 2000). Thus, to measure the level of shared understanding, we hypothesize that *the cohesiveness of communication in design teams is a measure of the level of shared understanding and commitment of the team*.

Given this assertion, we postulate that *teams with a high level of shared understanding will exhibit both design process quality and design outcome quality*. Figure 1 illustrates the relationship between document coherence, team cohesiveness, shared understanding, and process and design quality outcome.

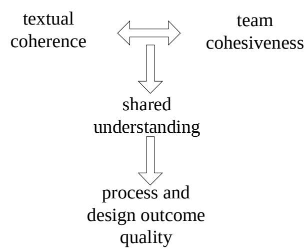

Figure 1. Correlation Between Document Coherence and Shared Understanding

A distinction is made between process quality, which relates to how a team performs as a team, and outcome, which relates to the quality of the designed artifact the team produced. To measure the level of shared understanding, we developed a technique for investigating textual coherence using latent semantic analysis. We cross-validated our technique for measuring shared understanding with expert reviews of the design teams, looking for correlations between our measurements of high shared understanding (as exhibited by high textual coherence) and the experts' assessment of high quality team process and design outcome.

We propose two methods to measure the cohesiveness of team communication. In the first method, we analyze the variance of the documents, the presumption being that shared understanding and variance have an inverse correlation. The second method analyzes the coherence of the documentation. The presumption is that coherence and shared understanding have a direct correlation and that teams with a strong shared understanding will distinguish themselves from other teams. Analytically, the variance measures the correlation *among* a team's documents whereas coherence measures the distinctiveness *between* one team's entire set of documentation and that of another team.

2. Methodology

2.1 GENERAL METHODOLOGY

The general approach to this study involves the capture of design documents such as mission statements, customer surveys, concept descriptions, concept selection rationale, prototype description and test plans, and design evaluation, generated within a design team working on a collaborative design project. Applying latent semantic analysis to the corpus of collected text requires computing singular vectors to model each document.

Correlation analysis of the singular vectors reveals attributes of the teams' communication patterns. The results of the correlation analysis will be compared to faculty and professional design judges' reviews of the design teams' process and outcomes.

2.2 TEXT PROCESSING

The teams' design documents as described in Section 1.1 were stored in a database. A word-by-document matrix, an example of which is shown in TABLE 1, was generated from this corpus by extracting key words from the natural language text. This word-by-document matrix F is comprised of *n* words w_1, w_2, …, w_n in m documents d_1, d_2, … d_m , where the weights indicate the total frequency of occurrence of term w_p in document d_q.

	d_1	d_2	…	d_m
w_1	0	1		2
w_2	1	0		0
…				
w_n	0	1		1

TABLE 1. Word by Document Matrix

We then convert the frequency counts to a log-entropy weighted word document matrix *X* using Equation 1.

$$X_{pq} = \frac{\log(F_{pq}+1)}{-\sum_{1-m}\left(\left(\frac{F_{pq}}{\sum_{1-m} F_{pq}}\right) * \log\left(\frac{F_{pq}}{\sum_{1-m} F_{pq}}\right)\right)} \tag{1}$$

where p=1…n and q=1…m.

2.3 LATENT SEMANTIC ANALYSIS

Latent semantic analysis (LSA) (Landauer et al., 1998) is a text analysis method that extracts the meaning of documents and of words by applying a statistical analysis to a large corpus of text to detect semantic similarities. The mathematical foundation for LSA lies in singular value decomposition (SVD), a method for reducing the dimensions of a matrix. The baseline theory for LSA is that by looking at the entire range of words chosen in a wide variety of texts, patterns will emerge in terms of word choice as well as word and document meaning. LSA is unique in its method of analyzing text; there is no consideration of word order or syntax. LSA was chosen as an analysis tool because of its demonstrated successes in identifying contextual meanings of documents, identifying the voice of a given document's author, and for analyzing the cognitive processes underlying communication (Landauer, 1999). Because identifying consensus and shared understanding in a corpus of communication requires more than simply correlating documents through key word similarity, LSA is an important computational tool to explore theoretically and empirically the links between meaningfulness of words and their social and intentional existence. LSA tells us that the higher-order associations contained in the usage of words generate social phenomena, such as "self", "personal identity" and "social group perception" (Landauer, 1999).

It is important to note that we utilized latent semantic analysis to measure textual coherence – not textual cohesiveness. Cohesiveness in text (lexical cohesion) is based on the pragmatics of grammar, or "the ties that bind a text together" (Halliday and Hasan, 1976) whereas textual coherence relates to the thematic consistency. That is, a coherent text presents (a) unified concept(s) "about the speaker's or author's purpose" (Fillmore, 1974). One recent study has demonstrated a structural link between lexical cohesion and textual coherence (Harabagiu, 1999).

As an example of the contrast between textual coherence and cohesiveness, consider the two excerpts from design documents by Team 5. The first (A) comes from the original project proposal by a team member, whereas the second (B) is the product description from the second draft of the product mission statement.

A: "There are very many different types of products that come under the description of roof racks and most car owners would admit to owning at least one system."

B: To provide a stylish, secure, reconfigurable storage and organizational platform for passenger cars, SUV's, wagons, and minivans, allowing active, recreational users fast and easy access to stuff.

In this example, the A is written in passive prose, whereas B is written as a directive. The sentences also utilized different grammatical constructs. Even though there is weak evidence of textual cohesiveness, there appears to exist a consistency of theme, storing and organizing items in cars. In fact, if all the documents authored by a team were collapsed into one, the grammar and structure of the various authors would likely not be textually cohesive. We used LSA to measure textual coherence, which we regard as a measure of shared understanding when that text arises from intra-team communication and design documentation.

Also, LSA cannot detect prosody in textual communication, that is, tone and rhythm. Prosody is an important device for accomplishing cohesion in spoken interaction that in writing is exhibited through the use of certain grammatical constructs and punctuation (Gumperz et al., 1984). Nor could LSA pick up on argumentation in text. If the text prefaced with terms such as “disagree” or “counter point,” yet dealt with a single concept or theme, our method based on LSA would measure an equivalent textual coherence to a document that prefaced text with terms such as “complete agreement” and “on the same page.” Yet, one would hardly characterize a team authoring the former documents as cohesive. Other techniques would need to be applied to detect argumentation (Raccah, 1995).

Singular value decomposition is applied to the log-entropy weighted word-by-document matrix *X* shown in Equation 1 resulting in three component matrices using the relationship $X = SVD^T$ where S is the n x m matrix of left *singular vectors*. The matrix dimension will typically be in the thousands to tens of thousands. However, the useful feature of LSA, in particular to SVD, is that one is able to discount all but about the first 300 singular values in each singular vector without losing much descriptive information. One aspect in applying LSA is determining the sufficient number of dimensions to retain to adequately capture context while reducing the number of retained dimensions to the smallest number possible in order to minimize computational complexity. The exact number of singular vectors to retain is subject to change as part of the algorithm tuning. These singular vectors s_1, s_2, … s_k are the basis for the remaining computations. TABLE 2 describes which singular vectors are retained for analysis in each Case A through J.

Case Descriptions	Singular Values Included in Analysis
Case A	Singular Values 1-10
Case B	Singular Values 2-11
Case C	Singular Values 1-30
Case D	Singular Values 2-31
Case E	Singular Values 1-100
Case F	Singular Values 2-101
Case G	Singular Values 1-200
Case H	Singular Values 2-201
Case I	Singular Values 1-300
Case J	Singular Values 2-301

TABLE 2. Experimental Case Descriptions

2.4 METHOD ONE: VARIANCE OF DOCUMENTS

To quantify shared understanding in design teams, we build upon research documented by Hill (Hill et al., 2001). Hill successfully applied latent semantic analysis upon a corpus of design documents to identify topical and voice similarities in design documents. These similarities, when combined, allow for the identification of a shared understanding. In order to identify teams that share a vision, we utilize latent semantic analysis and calculate the *s-dimensional* radius of a set of design documents generated by a given design team, where *s* is the number of singular values being retained for the given analysis. Because shared understanding – comprised of topical and voice similarity – is a self-defined signature of a team, this method identifies the level of a team's shared understanding by comparing the magnitudes and deviations of the s-dimensional radius of the documents in the LSA space. Conceptually, the centroid of the documents represents the shared understanding; dilution of that shared understanding would manifest through the variance of the radii and distance to outlying documents.

Based on this, our algorithm proceeds as follows:

1) Define *t* matrices m_i, i=1…t, denoting the total number of unique teams represented in the original word document matrix. Place all documents n_i generated by team *i* in the matrix m_i. Thus, the matrix m_i has v_i columns, the number of singular values, and n_i rows, the number of documents.
2) Calculate the centroid c_i of each team's document set by finding the average value of each column of m_i.

$$c_i = \frac{\sum_{k=1}^{v_i} s_{k,i}}{n_i} \tag{2}$$

where $s_{k,i}$ is the singular vector k representing a document produced by team i.

3) Calculate maximum (max), mean (μ) and standard deviation (σ) of the *s-dimensional* radius for each team.

$$\max{}_i = \max_{n_i}\left[\sqrt{\sum_{k=1}^{v} (s_{k,i} - c_i)}^{2}\right] \tag{3}$$

$$\mu_i = \frac{\sqrt{\sum_{k=1}^{v_i} (s_{k,i} - c_i)^2}}{n_i} \tag{4}$$

$$\sigma_i = \frac{\sqrt{\sum_{n_i}\left(\left(s_{k,i} - c_i\right) - \mu_i\right)^2}}{n_i} \tag{5}$$

We expect that teams that exhibit "good" shared understanding will have a smaller standard deviation and a lower maximum relative to the value of the mean of teams that exhibit "poor" shared understanding.

2.5 METHOD TWO: TEXTUAL COHERENCE

To quantify textual coherence, we build upon the work of Dhillon and Mohda (1999) who developed a technique for clustering very large sets of documents using LSA. Dhillon and Mohda determined that the coherence of a set of documents is simply a measure of the L^2 norm of the document vectors in the set. Analytically, the norm determines whether all singular vectors for a team's set of documents are identical. In other words, it is as if all the authors were writing expressing the same sentiment but in slightly different styles. We hypothesize that the coherence of a set of design documents is indicative of the cohesion of the design team. The idea is that cohesiveness of the team will be expressed in the generated documents as coherence. To do this analysis, the set of singular vectors generated from the documents of each team are grouped and the coherence is measured. The following method was employed.

1) Define t matrices m_i, i=1…t, denoting the total number of unique teams represented in the original word document matrix. Place all documents n_i generated by team i in the matrix m_i. As with the variance of documents, the matrix m_i has v_i columns and n_i rows.
2) Calculate the cohesiveness χ of each group, which according to Dhillon and Modha, is equivalent to computing the L^2 norm of the sum of the singular vectors representing the document set for each team.

$$\chi = \left\| \sum_{n_i} s_i \right\| \tag{6}$$

2.6 QUALITATIVE COMPARISONS

To validate our quantitative metrics, qualitative measures of the process and outcome quality for each team were taken into consideration. Each team is assigned a Rank from 1 to 9, 1 being the best and 9 the worst. The Rank is based on the cumulative assessments of ten professional product designers on the quality of the team's final product (in terms of the final product satisfying the design objectives as stated by the mission statement) and of the team's process quality (i.e., how well the team executed the product development process and functioned as a team). There were seven criteria: mission statement, customer and user needs, concept generation, concept selection, prototype feedback, financial analysis and final prototype. For each criterion, the judges assigned a score from 1 (worst) to 5 (best) based on recommended rubrics provided by the faculty. For example, in assessing a team's mission statement, the judge would examine whether the mission statement communicated the intent of the students' project, the clarity of the mission statement, and the description of the target market and business goals of the product. The judges based their assessments on 10-minute formal presentation by the team, a notebook recording the process undertaken, the final prototype, and informal discussions between the judges and each team during a three-hour "trade show." We then correlated the quantitative measures with the qualitative assessments.

To assure validity of the quantitative measure, we followed a protocol that maintained the anonymity of the teams and of the team members and removed potential bias in the data that might result from the teams' knowledge of the purpose of the study. First, we removed all the names of the teams and the names of the authors of the documents from the analyses. Second, the results of these analyses did not figure into the grading of the student teams. The instructors and judges had no prior knowledge of the analysis results. These practices ensured that the teams did not "massage" the collected data in an attempt to improve their grading.

3. Experiment and Results

3.1 DATA SET

Each design team submitted 18 design documents as part of their product development deliverables. A total of 135 documents were submitted. Each group turned in between 13 and 18 documents because some teams combined deliverables into one document. The documents types ranged from mission statements to financial analysis of the project to customer and user feedback summaries. The activities of the groups are summarized in TABLE 3. TABLE 3 shows how many design documents were generated, as well as how many team members contributed a significant number of documents. While the contributing member statistic is not a direct reflection of how the work was distributed among the team – in fact, it is only a measure of who turned the document in – it does provide insight into how the team distributed tasks.

Team	1	2	3	4	5	6	7	8	9
No. Documents Submitted	14	15	17	13	15	18	16	14	13
No. Contributing Members	2	2	5	3	5	4	5	4	3

TABLE 3. Summary of Group Activity

3.2 QUALITATIVE MEASURES

The results of the judging are shown in TABLE 4. All judges' ratings were included to establish the ranking. The teams are ranked in order from highest 1, to lowest, 9. In TABLE 4 and as in all subsequent tables, the top five performing teams are indicated in **bold**.

Team	**1**	2	**3**	4	**5**	6	**7**	**8**	9
Rank	**5**	7	**3**	9	**1**	6	**2**	**4**	8

TABLE 4. Team Ranking

The results of our quantitative analyses for shared understanding and cohesiveness were compared to the judges' assessments by identifying how each team performed relative to the median value of the metric. For example, in method one, we calculated the standard deviation of the s-dimensional radius for each team. We then calculated the median value. Those teams that performed better than or equal to the median are indicated with a '+' in the table of results, while those that performed worse than the median are indicated with a '–'. If we had perfect correlation between the judges assessment and the measurements, then the top 5 rated teams would show a plus and the bottom four would show a negative. Based on the correlation, we hope to draw some conclusions about the correlation between shared understanding and cohesiveness with process and design quality outcomes. More importantly, correlation between the quantitative measures and the judges' assessments would offer an initial basis for validating our methodology.

3.3 RESULTS OF METHOD ONE

In order to analyze the data, we generated tables summarizing each experimental case. The results of the document variation analyses (method one) are summarized in TABLE 5, TABLE 6, and TABLE 7.

Mean Distance Median	Case A	Case B	Case C	Case D	Case E	Case F	Case G	Case H	Case I	Case J	Average
Group 1	+	+	+	+	+	+	+	+	-	-	+
Group 2	+	-	+	+	+	+	-	-	-	-	-
Group 3	-	-	-	-	+	+	-	-	-	-	-
Group 4	-	-	-	-	-	-	-	-	-	-	-
Group 5	-	-	-	+	+	+	-	-	-	-	-
Group 6	+	+	+	+	-	-	+	+	+	+	+
Group 7	-	-	-	-	-	-	+	+	+	+	+
Group 8	-	+	-	-	-	-	+	+	+	+	-
Group 9	+	+	+	-	-	-	+	+	+	+	+

TABLE 5. Mean Euclidian Distance in LSA Space

Max Distance Median	Case A	Case B	Case C	Case D	Case E	Case F	Case G	Case H	Case I	Case J	Average
Group 1	+	+	+	+	+	+	-	-	-	-	+
Group 2	+	+	+	+	+	+	-	-	-	-	+
Group 3	-	-	-	-	+	+	+	+	-	-	+
Group 4	+	-	+	+	-	-	-	-	-	-	-
Group 5	+	+	+	+	+	+	-	-	-	+	+
Group 6	-	-	-	-	-	-	-	-	+	-	-
Group 7	-	-	-	-	-	-	+	+	+	+	-
Group 8	-	-	-	-	+	+	+	+	+	+	-
Group 9	-	+	-	-	-	-	+	+	+	+	-

TABLE 6. Maximum Euclidian Distance in LSA Space

St. Dev Distance Median	Case A	Case B	Case C	Case D	Case E	Case F	Case G	Case H	Case I	Case J	Average
Group 1	+	+	+	+	-	-	-	-	-	-	+
Group 2	+	+	+	+	+	+	-	-	-	-	+
Group 3	-	-	-	-	+	+	+	+	-	-	+
Group 4	+	+	+	+	-	-	-	-	-	-	-
Group 5	-	-	-	+	+	+	+	+	+	+	+
Group 6	-	-	-	-	-	-	-	-	+	+	-
Group 7	-	-	-	-	+	+	+	+	+	+	-
Group 8	+	-	+	-	+	+	+	+	+	+	+
Group 9	-	+	-	-	-	-	+	+	+	+	-

TABLE 7. Standard Deviation of Euclidian Distance in LSA Space

As can be seen in these three tables, the measure of increasing standard deviation correlates well with decreasing shared understanding. Cases E and F had the highest accuracy. They correctly identified Teams 3, 5, 7, 8 as having better than average shared understanding and Teams 4, 6 and 9 as having lower than average shared understanding. It mis-categorized Teams 1 and 2. Thus, on average, it had an accuracy of 78%.

Figure 2 illustrates the difference between levels of shared understanding. Team 3 created a superior product relative to Team 9 and had a superior process. Graphically, this manifests as greater dispersion in the documents (shown as 2-D LSA vectors) generated by Team 9 compared to Team 3. Since Cases E and F had the highest accuracy, we conclude that retaining the first 100 dimensions of the singular vectors (as shown in Table 2) is sufficient to capture the context of the documents. We will test this assertion again with method two.

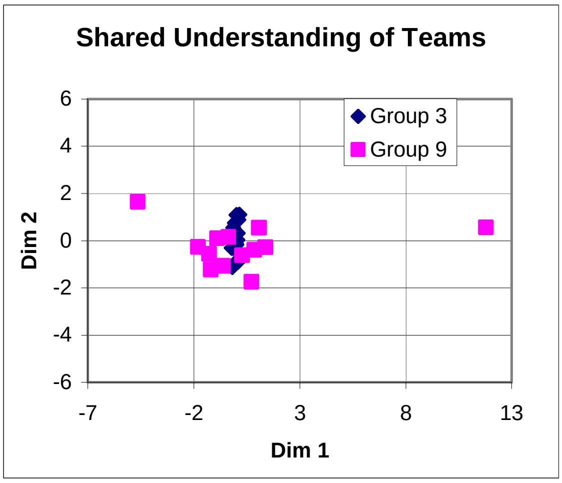

Figure 2. Comparison of Shared Understanding of Two Teams

The prior analysis tries to match the quantitative metrics from our studies to the qualitative scores given by judges. Examining the comments that judges made relative to high performing teams and low performing teams can have further insight into the analysis. The following are comments given to Team 3, a high performing team:

- "Direct, Specific and Clear"
- "Well thought out process to determine customer needs and generate concepts."
- "Good rigorous methods to organize all the features and combinations."

All told, the comments about Team 3, and about the high performing teams in general, indicated a positive impression of the team's vision and how the vision was executed. When contrasted with comments about low performing teams, as follows about Team 9, a clear difference can be seen in how this team's vision was created and executed. Comments about Team 9 are as follow:

- "Too broad. It would be difficult to screen potential product designs. Seems more like a mission statement for a broad range of products that provide similar services."
- "Focused too much on usability issues and not enough on functional requirements."
- "OK. But did not explore problem."

The sentiments reflected in the judges' comments were a lack of a clear vision, focusing on non-key issues and an approach that focused on a single solution too early in the process. These sentiments, we think, accurately describe a team without a shared vision.

3.4 RESULTS FROM METHOD TWO

In measuring the coherence of documents, we attempted to capture the cohesiveness of the design teams. As can be seen in TABLE 8, the measure of increasing coherence correlates well with increasing shared understanding. As in the above analysis, we use Cases E and F to calculate the accuracy. This method correctly identified Teams 1, 3, 7, and 8 as having better than average shared understanding and Teams 4, 6, and 9 as having lower than average shared understanding. It mis-categorized Teams 2 and 5. Thus, on average, it had an accuracy of 78%, equivalent to method one.

Coherence Median	Case A	Case B	Case C	Case D	Case E	Case F	Case G	Case H	Case I	Case J	Average
Group 1	-	-	+	+	+	+	+	+	+	+	-
Group 2	-	-	+	+	+	+	+	+	+	+	-
Group 3	-	-	-	-	+	+	+	+	+	+	-
Group 4	-	-	-	-	-	-	-	-	+	+	+
Group 5	+	+	-	-	-	-	+	+	-	-	+
Group 6	+	+	+	+	-	-	-	-	-	-	+
Group 7	+	+	+	+	+	+	-	-	+	+	+
Group 8	+	+	+	+	+	+	+	+	-	-	+
Group 9	+	+	-	-	-	-	-	-	-	-	-

TABLE 8. Cohesiveness

3.5 RESULTS SUMMARY

TABLE 9 summarizes the experimental cases that were run, showing the accuracy of each metric in sorting the teams according to the judges' assessments. Equation 7 defines accuracy.

$$Accuracy = \frac{TP + TN}{\sum g_i} \qquad (7)$$

TP represents the True Positives found in the study, that, is high quality teams as deemed by the judges is also identified by the metrics and TN represents the True Negatives, that is, low quality. The denominator represents the total number of teams being studied, in this study nine. As can be seen in both TABLE 9 and Figure 3, the standard deviation and coherence metrics turn out to be the best method for measuring the shared understanding of design teams.

Metric Summary	Case A	Case B	Case C	Case D	Case E	Case F	Case G	Case H	Case I	Case J
Coherence	**56%**	**56%**	**56%**	**56%**	**78%**	**78%**	**78%**	**78%**	**56%**	**56%**
Mean Dist.	**22%**	**44%**	**22%**	**44%**	**67%**	**67%**	**56%**	**56%**	**44%**	**44%**
Max Dist.	**33%**	**33%**	**44%**	**44%**	**67%**	**78%**	**67%**	**67%**	**44%**	**56%**
St. Dev Dist.	**44%**	**22%**	**44%**	**44%**	**78%**	**78%**	**78%**	**78%**	**56%**	**56%**

TABLE 9. Accuracy of Metrics

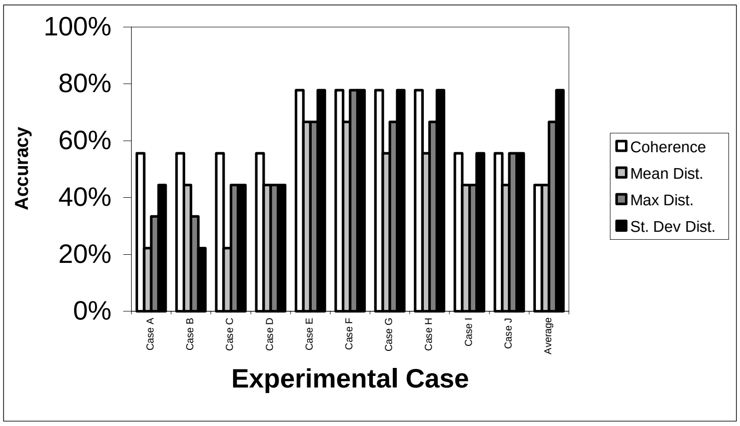

Figure 3. Summary of Evaluation of Metrics

As stated in Section 1.2, our premise is that teams with a high level of shared understanding will exhibit both design process quality and design outcome quality. Using our methodology to mine design documents for levels of shared understanding based on variance in communication and textual coherence, we were able to distinguish between teams with high and low shared understanding. We cross-correlated our quantitative measurements to judges' assessments of the teams. We found up to 80% agreement between our quantitative measures and the judges' assessments. Because we have shown that our method can find teams with high levels of shared understanding, and that these same teams had high quality outcomes based on impartial judges' assessments, then, based on the premise described in Figure 1, we have made progress towards demonstrating the validity of our methodology for measuring levels of shared understanding in design teams.

4. Discussion and Implications

This paper addresses the question of identifying the level of shared understanding in design teams using the computational linguistic tool of latent semantic analysis. By computing the standard deviation (variance) of the s-dimensional radius of the singular vectors representing the design team's documents and the textual coherence of the documents, we were able to sort the design teams in order of level of shared understanding. We correlated our metrics of shared understanding to judges' assessments of the process quality and design quality of the design teams. We also found a correlation between textual coherence and team cohesiveness. Finally, we found some quantitative support to suggest that a high level of shared understanding and cohesiveness among the design team leads to a better process and design quality outcome. As such, we conclude that, with about 80% agreement to expert human assessments in our study, our methodology can distinguish between levels of shared understanding in design teams based on the team's design documentation and that these levels correlate with increasing process and outcome quality.

The validation of our methodology provides exciting possibilities in future implementations. First, the experimental results from applying the metrics of variance and cohesion to the document show that retaining the 100 most significant singular vectors is sufficient to capture the context of the documents while minimizing the computational complexity. Also, the 100 most significant singular vectors provide the most accurate results. Future studies of design team communication using our methodology can use this guideline for retaining the singular values. Second, we were able to differentiate high and low performing teams with good correlation to expert assessments.

As a tool for reflective teams, they can use our methodology to reflect on their progress through the design process and identify potential problems in their performance in a predictive manner. Because most design teams use a document management or product data management system to help manage the design project, the technique can be run over the design documentation as each engineer contributes new documents. Each team can benchmark itself in real-time and use information on team cohesion to address deficiencies when appropriate. However, not all design documentation would be appropriate for this type of analysis. For example, documents from a finite element analysis of a component or structured data from a Taguchi quality test would not capture any meaningful dialog between the designers that would indicate their level of shared understanding. Minutes from design team meetings, studio reviews of the design, formal stage-gate reviews, product mission statements and other similar documentation which capture the thoughts and contain a “dialogicality of communication” between designers would be suitable candidates for this type of analysis.

This research offers significant advances in studying the communication of design teams to ascertain the level of shared understanding and cohesiveness. As a possible tool for reflection, design teams might apply these metrics to measure how well they compare respect to other design teams. In other words, they can use it as an indicator of the shared understanding and cohesiveness of their own design teams, taking steps as necessary to improve these psychosocial metrics as desired. For example, product development teams could use it for personal reflection and to measure the level of commitment of their teammates to adopt appropriate strategies of persuasion. We believe that this could present a useful addition to computer-support collaborative work tools that support collaborative design teams.

While this research offers insight into shared understanding based on the final design documentation of the design teams, we believe that this area has much more potential for investigation with respect to communication in real-time, such as via e-mail and chat software. Further, by studying communication as it happens and through various stages of design, we can study the change in shared understanding to detect if there are patterns of shared understanding that lead to better design outcomes. For example, are design teams that reach shared understanding too quickly and never diverge less likely to produce novel designs? Finally, what is the implication of the mis-categorized teams? We postulate that while a high level of shared understanding and cohesiveness increases the likelihood of a high quality process and outcome, they are not guarantees.

The changing nature of teams from small, closely situated teams to larger, widely dispersed teams whose communication is nearly exclusively via information technology tools such as e-mail, places people in an information space that is not only too large for any single person to know completely, but that is also growing and changing at an increasing rate. Bringing tools to assist teams in gauging their communication may ultimately help them to improve team performance. We hope this line of research will have potential application to a range of enterprises that depend on effective teams for success. Finally, the method of latent semantic analysis shows great promise as a tool for modeling and evaluating the cognitive and psychosocial behavior of design teams.

5. Acknowledgements

The authors wish to acknowledge the support of the faculty and the students of the new product development course at UC Berkeley in providing a test-bed for this research.

6. References

Adams, BN: 1967, Interaction Theory and the Social Network, *Sociometry*, **30**(1): 64-78.

Baird, F, Moore, CJ, and Jagodzinski, AP: 2000, An ethnographic study of engineering design teams at Rolls-Royce Aerospace, *Design Studies*, **21**(4): 333-355.

Bødker, K and Pedersen, JS: 1991, Workplace Cultures: Looking at Artifacts, Symbols and Practices, In *Design at Work: Cooperative Design of Computer Systems*, J. Greenbaum, M. Kyng (eds.), Lawrence Earlbaum Associates Inc., Hillsdale, New Jersey, 1991, pp. 121-136.

Bucciarelli, LL: 1994, *Designing Engineers,* MIT Press, Cambridge MA.

Citera, M, McNeese, MD, Brown, CE, Selvaraj, JA, Zaff, BS, and Whitaker, RD: 1995, Fitting Information Systems to Collaborating Design Teams, *Journal of the American Society for Information Science,* **46**(7): 551-559.

Cook KS, and Whitmeyer, J.: 1992, Two Approaches to Social Structure: Exchange Theory and Network Analysis, *Annual Review of Sociology*, **18**:109-127.

Cooper, RG, and Kleinschmidt, EJ: 1995, Benchmarking the firm's critical success factors in new product development, *Journal of Product Innovation Management*, **12**(5), November 1995: 374-391.

Dhillon, IS, and Modha, DS: 1999, Concept Decompositions for Large Sparse Text Data Using Clustering, IBM Almaden Research Center, Research Report RJ 10147 (95022), July 8, 1999.

Dorst, K: 1997, *Describing design – a comparison of paradigms*, Doctoral Thesis, Delft Technical University, The Netherlands.

Griffin, A: 1995, PDMA Research on New Product Development Practices: Updating Trends and Becnhmarking Best Practices, *The Journal of Product Innovation Management,* **14**(6): 429-458.

Fillmore, CJ: 1974, The future of semantics, *Berkeley Studies in Syntax and Semantics*, Department of Linguistics and Institute of Human Learning, University of California, Berkeley, CA, pp. 1V-5.

Gumperz, JJ, Kaltman, H, and O'Connor, MC: 1984, Cohesion in Spoken and Written Discourse: Ethnic Style and the Transition to Literacy, *Coherence in Spoken and Written Discourse*, D. Tannen, (ed.), ABLEX Publishing Corporation, Norwood, New Jersey, pp. 3-19.

Halliday, MAK and Hasan, R: 1976, *Cohesion in English*, Longman, London.

Harabagiu, S: 1999, "From lexical cohesion to textual coherence: A data driven perspective," *International Journal of Pattern Recognition and Artificial Intelligence*, **13**(2): 247-265.

Henke, JW, Krachenberg, AR, and Lyons, TF: 1993, Cross-Functional Teams; Good Concept, Poor Implementation, *The Journal of Product Innovation Management,* **10**(3): 216-229.

Hill, A, Song, S, Dong, A and Agogino, AM: 2001, Identifying Shared Understanding in Design Using Document Analysis, *Proceedings of the 2001 ASME Design Engineering Technical Conferences*, September 9-12, 2001, Pittsburgh, PA, DETC2001/DTM-21713.

Landuaer, TK: 1999, Latent semantic analysis: A theory of the psychology of language and mind, *Discourse Processes*, **27**(3): 303-310.

Landauer, TK, Foltz, PW and Laham, D: 1998, Introduction to Latent Semantic Analysis, *Discourse Processes*, **25**: 259-284.

Lloyd, P: 2000, Storytelling and the development of discourse in the engineering design process, *Design Studies*, **21**: 357-373.

McDonough, III, EF, Kahn, KB, and Barczak, G: 2000, An investigation of the use of global, virtual, and collocated new product development teams, *The Journal of Product Innovation Management*, **18**(2): 110-120.

Moody, J and White, D: 2000, Social Cohesion and Embeddedness: A hierarchical conception of social groups, Santa Fe Institute, Working Paper No. 00-08-049.

Neck, CP and Manz, CC.: August 1994, From groupthink to teamthink: Toward the creation of constructive thought patterns in self-managing work teams, *Human Relations*, **47**(8): 929-952.

Pahl, G and Beitz, W: 1996, *Engineering Design*, Second Edition, Springer-Verlag, London.

Gumperz, JJ, Kaltman, H, and O'Connor, MC: 1984, Cohesion in Spoken and Written Discourse: Ethnic Syle and Transition to Literacy, *Coherence in Spoken and Written Discourse*, D. Tannen, (ed.), ABLEX Publishing Corporation, New Jersey, pp. 3-19.

Raccah, PY: July 1995, Argumentation and Natural Language - Presentation and Discussion of Four Foundational Hypotheses, *Journal of Pragmatics*, **24**(1-2): 1-15.

Robinson, M and Bannon, L: 1991, Questioning Representations, *Proceedings of ECSCW'91: 2nd European Conference on Computer-Supported Collaborative Work*, L. Bannon, M. Robinson, and K. Schmidt, (eds.), Kluwer, Amsterdam Sept 25-27, 1991, pp. 219-233.

Schön, DA: 1983, *The reflective practitioner: how professionals think in action*, Basic Books, New York.

Simon, HA: 1996, *The sciences of the artificial*, Third Edition, MIT Press, Cambridge.

Valkenburg, R: 2000, *The Reflective Practice in product design teams*, Doctoral Thesis, Delft Technical University, The Netherlands.

Wasserman, S and Galaskiewicz, J (eds.): *Advances in social network analysis: research in the social and behavioral sciences*, Sage Publications, Thousand Oaks, CA.