It gives me great pleasure to view the web publication of the online research journal of The Graduate Student Advisory Committee (GSAC), *The Hilltop Review: a Journal of Western Michigan University Graduate Research*. The journal provides a venue to share the outstanding scholarly and creative activities of the graduate students at Western Michigan University. This peer-reviewed journal contains manuscripts that represent the disciplines of Philosophy, Paper Engineering and Chemical Engineering and Imaging, Educational Studies, Communications, and Mechanical and Aeronautical Engineering. In addition it displays a painting by a student in the Master of Fine Arts program. Students from across the campus have come together to contribute to the academic life of the University.

During the five years that I was involved within the Graduate College at WMU, I watched GSAC mature and grow into a vibrant organization that represents the needs of graduate students. GSAC has taken a leadership role in advocating for graduate study, contributing to the University community, and establishing projects that benefit the Kalamazoo community. Now with the addition of *The Hilltop Review* GSAC has expanded its scope to share the academic and cultural contributions of its members. It has created a journal that will highlight both the research rigor and creative talents of graduate students at WMU. Congratulations GSAC! Job well done.

Bill Wiener, Vice Provost and Dean
Marquette University
Dear Friends and Colleagues,

I am excited to welcome you to the first edition of *The Hilltop Review: A Journal of Western Michigan University Graduate Research*. I would first like to congratulate the people whose submissions were accepted into the journal. Thank you for making the first edition of this publication diverse and brimming with knowledge.

Next, I would like to thank someone who offered the journal a different type of submission. Andrew McNair was gracious enough to let the committee use his amazing artwork for the cover of the journal. Most importantly, I would like to thank my committee for doing so much to help this journal become a reality. It would not have happened without the hard work of Todd Ide, Jennifer Smith, and Jason Olsen.

Finally, I would like to thank Dr. William Wiener and Dr. Kevin Vichcales from The Graduate College for their guidance throughout the process of producing the journal. Thanks again to everyone. I hope you all enjoy *The Hilltop Review* for many years to come!

Toni N. Strutz

Editor, *The Hilltop Review*
Competition and Interpersonal Conflict in Same-Sex Platonic Friendships

EWA URBAN
School of Communication
Western Michigan University

ABSTRACT

This research explores the connection between competition and distributive and integrative conflict messages in same-sex platonic friendships. The study is a partial replication of Messman and Mikesell’s (2000) examination of competition and interpersonal conflict in dating relationships. The competitive behaviors of same-sex friends examined here are compared with those of dating couples as explored by Messman and Mikesell. Findings of the current research indicate that increased competition between same-sex friends is linked to a rise in distributive conflict tactics as well as to a reduction in integrative messages, which is consistent with Messman and Mikesell’s results. The current study also investigates the similarities and differences between men and women in their approach to competition while interacting with their same-sex platonic friends. In this study, significant differences between men’s and women’s competitive behaviors were found. This result contrasts with the findings of Messman and Mikesell, who concluded that no distinctions were apparent between males and females in romantic relationships in regard to their competitive interactions.

INTRODUCTION

Competition is one of the intrinsic characteristics of mainstream culture in the United States (Aronson, 1999; Tjosvold, Johnson, Johnson, & Sun, 2003). The American society was founded on an ideal of each individual being self-reliant and struggling to meet his or her desires and needs while attempting to outperform others. Highly individualistic Americans tend to constantly compete against their own weaknesses and the strengths of others in order to develop their full potential, prove their superiority, differentiate themselves, and acquire high status (Hirschman, 2003). Doing well in American culture frequently corresponds to beating someone else (Aronson, 1999). This society of winners and losers avails itself of a language that mirrors these values. In a culture where one “wins a promotion, beats the other sales clerks, outsmarts a teacher, becomes a superstar, defeats enemies, and is the best student,” a day without competition is rather unlikely (Tjosvold et al., 2003, p. 63). It permeates all aspects of America including the economy, education, leisure, and the workplace. Thus, it is unavoidable in interpersonal relationships (Kohn, 1986).

In order to test the supposition that individuals in U.S. society compete while interacting with one another, empirical study of competition in interpersonal relationships is indispensable. The present study explores the connection between competition and conflict interaction in same-sex platonic friendships and examines the similarities and differences between men and women in their approach to competition as it relates to their friendships. This study is a partial replication of the research conducted by Messman and Mikesell (2000), who investigate the nature of competition among men and women in dating relationships and the link between competition and conflict strategies.

In the following section, the connection between competition and conflict, as well as varying stances on competition in interpersonal relationships, will be reviewed. Next, the findings concerning men’s and women’s competitive behavior and conflict interaction strategies will be presented. Finally, the nature of the interpersonal relationships investigated in the study will be scrutinized. Research on men’s and women’s same-sex friendships will be explored in order to illustrate an array of views on the characteristics of these relationships.

COMPETITION, CONFLICT MESSAGES AND SEX DIFFERENCES

Competition and Conflict

Numerous researchers consider competition to be one of conflict management strategies or styles, and contrast it with cooperation (Canary, Cunningham, & Cody, 1988; Cupach & Canary, 2000; Kilmann & Thomas, 1977;
Competition and Interpersonal Conflict

Sillars, 1980). According to Kilmann and Thomas, a competing style is utilized by individuals who desire to defeat the other party in an attempt to meet their own concerns while ignoring the needs of others. Sillars argues that competitive or distributive conflict strategies decrease the likelihood of successful conflict resolution because they generate competitive attributions about the partner. He characterizes distributive strategies as those that refer to attempts at conflict resolution that involve concession seeking, individualistic goal orientation, and a negative evaluation of the other. Integrative strategies, on the other hand, promote the exchange of information, open conversation about conflict, and positive or neutral perception of the interactants.

The researchers’ perceptions of competition as well as its impact on interpersonal relationships vary. For some, it is inextricably interwoven with distributive conflict behaviors and relational ruin (Deutsch, 1973; Kohn, 1986). For others, competition does not appear to be inherently destructive. In fact, it can benefit individuals because of its positive effect on self-esteem (Meeker, 1990). It can serve as a healthy catalyst for relationships, and is not restricted to distributive conflict tactics (Cheng & Chan, 1999; Hartup, 1992; Rawlins, 2001; Tjosvold et al., 2003).

The present study investigates the relationship between the areas of competition revealed by the research conducted by Messman and Mikesell (2000) and integrative and distributive conflict strategies in same-sex platonic friendships. Specifically, the research question posed is:

**RQ 1:** How are areas of competition associated with integrative and distributive conflict behaviors in same-sex platonic friendships?

**Women, Men and Competitive Behaviors**

The relationship between gender and competition has been explored by a plethora of researchers (Alagna, 1982; Benenson, Roy, & Waite, 2002; Campbell, 1999; Campbell, 2004; Meara & Day, 1993; Meeker, 1990; Robinson & Lipman-Blumen, 2003; Staley & Cohen, 1988). Their findings as to who tends to be more competitive in their interpersonal interactions, however, are inconclusive. Several authors argue that males are likely to behave more competitively than females (Alagna, 1982; Kohn, 1986; Platow & Shave, 1995; Staley & Cohen, 1988) while others claim the opposite (Meara & Day, 1993; Spitzberg, 1997). Yet, a number of studies demonstrate that there are no sex differences in the desire to compete (Benenson et al., 2002; Campbell, 1999; Campbell, 2004; Meeker, 1990) and that individual’s behavior cannot be predicted on the basis of their biological sex (Cupach & Canary, 1995).

For some scholars, the concept of competitiveness is the feature that separates femininity from masculinity (Bem, 1974; Richmond & Martin, 1998). The stereotypical assumption that men tend to be more competitive, aggressive and forceful is confirmed by Staley and Cohen’s (1988) study. Alagna (1982) suggests that women do tend to compete less than men. She explains her findings in terms of women’s greater sensitivity to being evaluated by their peers, who may perceive their competitive behaviors to be at variance with stereotypical sex role norms. They may strive for more positive interactions, avoid tension and strain, which they feel may endanger their relationships (Voss, Markiewicz, & Doyle, 1999; Wright, 1982).

Competition is by no means restricted to one sex. Although in the American culture, men are typically perceived as more assertive or masculine than women, whose communicative style is described as responsive or feminine, both sexes can display both characteristics, depending upon situational, relational, cultural and individual aspects (Bem, 1974; Cupach & Canary, 1995; Meeker, 1990; Richmond & Martin, 1998; Winstead, Derlega, & Rose, 1997). Men’s and women’s willingness to compete in specific situations can also be contingent upon the sex of the partner with whom they are interacting (Meara & Day, 1993). Since both sexes are subject to the same social and cultural influences that value competition, they are likely to equally engage in competitive behaviors (Meara & Day, 1993; Robinson & Lipman-Blumen, 2003). As the foregoing review of the literature indicates, however, the relationship between competition and gender still remains ambiguous. Messman and Mikesell (2000) explore competitiveness between heterosexual dating partners, whereas the present study examines sex differences and competitive behaviors in another type of interpersonal relationship, namely same-sex platonic friendships and poses the following question:

**RQ 2:** Do men and women differ in the frequency of their reports of areas of competition in same-sex platonic friendships?

**Same-Sex Platonic Friendships**
Friendships may vary in intensity, but in essence, they are all voluntary, preferential and gratifying relationships in which individuals are committed to each other (Hartup, 1992; Laursen, 1996; Wright, 1982; Wright, 1985). Trust, reciprocity and equity are inherent features of friendships (Rawlins, 2001; Winstead et al., 1997; Wright, 1982; Wright, 1985). All close friendships, regardless of the sex of individuals involved, entail the sharing of activities and interests, intimacy, emotional support, and self-disclosure (Wright, 1985). They are not divorced from tension and stress, however, since they may involve competing for rewards (Winstead et al., 1997; Wright, 1982; Wright, 1985).

Several researchers assert that women’s same-sex platonic friendships differ notably from men’s as to the type of the interaction, although the importance of the relationship is comparable for both sexes (Roy, Benenson, & Lilly, 2000). These differences have been associated with distinct socialization patterns (Winstead et al., 1997). Generally, men’s friendships involve joint activities, especially physical and competitive sports, while the core of women’s friendships consists of mutual disclosure and sharing, greater intimacy and closeness (Parker & de Vries, 1993; Rawlins, 2001; Roy et al.; Wright, 1982).

Wright (1998) challenges the stereotypical dichotomy of the nature of friendships for the two sexes. He contends that the characteristics of women’s friendships and men’s friendships are more fundamentally similar than different. Wright posits that the fact that men’s friendships are considered less communal than women’s friendships might be due to the distinct manner in which men express intimacy. Women create closeness through conversations and self-disclosure; however, talking is not the only way to show intimacy (Black, 2000; Rawlins, 2001). In fact, engaging in shared activities is an equally valid expression (Winstead et al., 1997; Wright, 1998). Although women are somewhat more inclined to treat their friendships more holistically than men, the differences between the two types of relationships are not fundamental (Wright, 1982; Wright, 1998). Moreover, the frequently cited distinct nature of men’s and women’s friendships becomes even less conspicuous with the increased closeness and longer involvement in the relationship (Wright, 1982).

Women report less conflict with their same-sex platonic friends than do men (Voss et al., 1999; Black, 2000). This could either be due to the different definitions of conflict that the two sexes might hold, or to the fact that women experience less conflict in their close same-sex relationships than men, perhaps because of the greater closeness and intimacy of their relationships (Voss et al., 1999). Direct competition with same-sex friends is likely to create more discomfort in females than in males (Benenson et al., 2002; Cheng & Chan, 1999). Female competition might be perceived more negatively by their same-sex peers, which could possibly disturb the relational equilibrium (Alagna, 1982; Benenson et al., 2002). Campbell (1999, 2004) asserts that females are less likely to compete directly with other females because competition is less beneficial to them biologically. This does not mean, however, that they do not compete at all with one another; instead, they are more inclined to select the forms of competition that are less conducive to physical injury.

Friendships and romantic relationships share several characteristics. They both entail affection, loyalty, substantial relational investment, trust and commitment (Rawlins, 2001). Hartup (1992) and Laursen (1996) posit that friendships, as compared with romantic relationships, involve less closeness between partners. Thus, conflict is less likely to lead to relational improvement among friends, since they are less concerned about the termination of the relationship. The question remains whether the conflict interaction behaviors and approach to competition are the same or different in same-sex platonic friendships as in dating relationships. Hence, this study attempts to compare the link between competition and integrative and distributive conflict strategies in same-sex platonic friendships with those in dating relationships as reported by Messman and Mikesell (2000). Thus, the third research question posed here is:

RQ 3: How do same-sex friendships and heterosexual dating couples differ in their conflict strategies and in the relationships between areas of competition?

METHODS

Procedures

As in the original research conducted by Messman and Mikesell (2000), the data for the current study have been collected from students at a large Midwestern university. The students were contacted during a lecture and
briefed on the nature of the research. They were informed that participation was voluntary, that they would receive extra credit for taking part in the study, and that their identities would remain confidential. Those who agreed completed the questionnaire within 15-20 minutes. The students were asked to respond to questionnaire items with a particular same-sex platonic friend in mind. They were requested to write the initials of this friend in order to ensure that they were thinking of a particular person as they were answering all the questions.

Participants

A total of 204 students participated in the study. All the questionnaires were suitable for analysis. Although in the original study by Messman and Mikesell (2000) the sample consisted of 449 students, the number of participants asked for input in the present research was considered sufficient for the analyses performed here.

In the current study, 132 participants were female (64.7%) and 72 were male (35.3%). Seventeen students were African American (8.3%), three were Asian American (1.5%), and 107 were European Americans (52.5%). There was one Hispanic American (0.5%) and one Native American student (0.5%). The remaining 75 students (36.8%) reported their ethnicity/race as other. More than half of the participants (55.4%) were between 18 and 20 years old. Almost forty-one percent of the students (40.7%) were between the ages of 21 and 25; 1.5% were between 25 and 30 years of age, and 2% were over 30. Students’ class standing was as follows: 3.9% were freshmen, 31.9% were sophomores, 38.2% were juniors, and 25% were seniors. There was one graduate student (0.5%), and one participant (0.5%) reported his or her class standing as other.

One hundred and twenty-two students (61.3%) responded to the questionnaire with their best friend in mind, 57 students (28.6%) thought about a close friend while answering the questions, 19 students (9.5%) identified their relationship as good friends, and one (0.5%) referred to his or her same-sex platonic friend as a casual friend. Five students did not respond to the “type of friend” item. The length of friendships reported varied, with 55.4% reporting involvement of longer than 5 years, 12.7% having been friends for 3-5 years, 26.5% reporting friendships lasting between 1 and 3 years, and 5.4% answering the questions in reference to a friend whom they have known for less than one year.

Instrumentation

Consistent with Messman and Mikesell’s (2000) methods, the questionnaire utilized for the present study was comprised of two parts. Competition was defined on the first page of the questionnaire as “attempting to out-do or keep up with your friend.” The first part of the survey included 21 questions that asked the participants to assess the frequency of competing with their same-sex platonic friend using a 5-point Likert scale (1 = Never, 2 = Seldom, 3 = Sometimes, 4 = Often, and 5 = Always). The 21 questionnaire items employed in the study were the same as those used by Messman and Mikesell. Three of the original questions were modified to improve their clarity. Specifically, the question: “Do you compete over physical appearance?” was changed to “Do you and your friend compete over who is more physically attractive?”; “Who has more relational skills?” was changed to “Who has better relational skills?”, and “Who gets better offers?” was modified to “Who makes more money?” The questions fell into one of the seven categories of competition identified by Messman and Mikesell: (1) Achievement, (2) Social Skills, (3) Altruism, (4) Social Attractiveness, (5) Controversy, (6) Affection, and (7) Play. The items were ordered randomly; however, care has been taken to place the more difficult questions between easier ones.

The second part of the questionnaire was preceded by the following definition of interpersonal conflict: “a significant struggle over incompatible goals between you and this particular same-sex friend.” The participants were reminded to keep the same relationship in mind and to think of a particular conflict with this person while responding to a set of 14 questions. They reported how much they agreed or disagreed that they used each of the 14 tactics (integrative or distributive) in a particular conflict with their same-sex platonic friend utilizing 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neither Disagree/Nor Agree, 4 = Agree, and 5 = Strongly Agree).

RESULTS

Research Question 1 referred to the association between the seven areas of competition identified by Messman and Mikesell (2000) and the two conflict strategies: integrative and distributive. In order to examine the relationship between competition and conflict strategies in same-sex platonic friendships, Pearson Correlations were computed. All the correlations between the seven categories of competition and the two conflict strategies are included in Table 1.
Distributive messages were positively and significantly correlated with all seven areas of competition, indicating that an increase in all these forms of competition is related to a rise in distributive conflict tactics (p ≤ .01). The range of variance accounted for was from 20% to 40%, with Play accounting for 20%, Social Skills 22%, Affection 25%, Controversy 32%, Altruism and Achievement 37%, and Social Attractiveness 40%.

Integrative messages were negatively correlated with all five categories of competition: Achievement, Altruism, Social Attractiveness, Controversy and Play, meaning that greater competition in these areas is linked to a reduction in integrative conflict tactics. The negative correlations between three areas of competition – Play, Achievement and Social Attractiveness – and integrative messages were significant, accounting for 20%, 23% and 28% of variance respectively. However, Altruism and Controversy, which accounted for 11% of variance, were negatively correlated with integrative messages but not at a significant level. There was no correlation between Social Skills and Affection and integrative messages.

Research Question 2 inquired whether the reports of the frequency in the seven areas of competition differ for men and women engaged in same-sex platonic friendships. Independent t-tests were computed to investigate the similarities and differences between the sexes on all of the categories. The results are reflected in Table 2.

Men reported to compete more than women in four areas: Achievement, Social Attractiveness, Controversy, and Play, and the difference between the sexes was significant in all of them (p ≤ .01). Women reported to be more competitive than men in two areas: Affection and Social Skills. There was no reported difference between males and females in one category of competition: Altruism.

As far as conflict messages utilized among same-sex friends, women stated that they used more integrative tactics (M = 3.96, SD = .72) than men did (M = 3.66, SD = .74), and this difference was significant (t = -2.72, df = 202, p = .00). There was no significant sex difference, however, in the usage of distributive tactics, with the mean for men being 2.64 (SD = .85), and for women: 2.49 (SD = .86; t = -1.21, df = 202, p = .22).

TABLE 1
Correlations between Categories of Competition and Conflict Strategies in Same-Sex Platonic Friendships

<table>
<thead>
<tr>
<th></th>
<th>Integrative</th>
<th>Distributive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement</td>
<td>-.23*</td>
<td>.37*</td>
</tr>
<tr>
<td>Social Skills</td>
<td>.00</td>
<td>.22*</td>
</tr>
<tr>
<td>Altruism</td>
<td>-.11</td>
<td>.37*</td>
</tr>
<tr>
<td>Social Attractiveness</td>
<td>-.28*</td>
<td>.40*</td>
</tr>
<tr>
<td>Controversy</td>
<td>-.11</td>
<td>.32*</td>
</tr>
<tr>
<td>Affection</td>
<td>.00</td>
<td>.25*</td>
</tr>
<tr>
<td>Play</td>
<td>-.20*</td>
<td>.20*</td>
</tr>
</tbody>
</table>

* p ≤ .01

TABLE 2
Means and Standard Deviations by Sex for Competition Categories and Conflict Strategies

<table>
<thead>
<tr>
<th></th>
<th>Men's Means (Standard Deviations)</th>
<th>Women's Means (Standard Deviations)</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement</td>
<td>2.33 (.76)</td>
<td>1.89 (.70)</td>
<td>4.20</td>
<td>202</td>
<td>.00*</td>
<td>.44</td>
</tr>
<tr>
<td>Social Attractiveness</td>
<td>2.36 (.91)</td>
<td>1.97 (.86)</td>
<td>2.95</td>
<td>202</td>
<td>.00*</td>
<td>.38</td>
</tr>
<tr>
<td>Controversy</td>
<td>2.27 (1.02)</td>
<td>1.93 (.88)</td>
<td>2.47</td>
<td>202</td>
<td>.01*</td>
<td>.33</td>
</tr>
<tr>
<td>Play</td>
<td>3.30 (.98)</td>
<td>2.07 (.88)</td>
<td>9.10</td>
<td>202</td>
<td>.00*</td>
<td>1.22</td>
</tr>
<tr>
<td>Affection</td>
<td>1.97 (1.04)</td>
<td>2.45 (1.14)</td>
<td>-2.92</td>
<td>200</td>
<td>.00*</td>
<td>-.47</td>
</tr>
<tr>
<td>Social Skills</td>
<td>1.77 (.77)</td>
<td>2.04 (.98)</td>
<td>-2.01</td>
<td>202</td>
<td>.04</td>
<td>-.27</td>
</tr>
<tr>
<td>Altruism</td>
<td>1.84 (.76)</td>
<td>1.95 (.82)</td>
<td>-9.5</td>
<td>202</td>
<td>.34</td>
<td>-.11</td>
</tr>
<tr>
<td>Integrative</td>
<td>3.66 (.74)</td>
<td>3.96 (.72)</td>
<td>-2.72</td>
<td>202</td>
<td>.00*</td>
<td>-.29</td>
</tr>
<tr>
<td>Distributive</td>
<td>2.64 (.85)</td>
<td>2.49 (.86)</td>
<td>1.21</td>
<td>202</td>
<td>.22</td>
<td>.15</td>
</tr>
</tbody>
</table>

* p < .01
Research Question 3 asked whether the relationships between the categories of competition and the two conflict strategies are the same or different in same-sex platonic friendships as in heterosexual dating couples. The aforementioned results of this study were compared with Messman and Mikesell’s (2000) findings. These comparisons are included in Table 3.

For both types of relationship, distributive messages were positively and significantly correlated with all categories of competition. For same-sex platonic friendships, the variance ranged from 20% to 40%, whereas for dating heterosexual partners, it was from 14% to 51%. These findings suggest that, for both same-sex friendships and romantic relationships, an increase in distributive conflict behavior is marked by a rise in competitiveness.

There were some differences between the two kinds of interpersonal relationships in the correlations of the integrative conflict tactics and the competition categories. For both same-sex friendships and romantic couples, the correlations were negative and significant in the areas of Social Attractiveness (accounting for 28% of variance for friends, and 17% for couples) and Achievement (23% of variance for friends, and 18% for romantic partners). For Controversy and Altruism categories, the negative correlations were at a significant level for couples (with 24% and 23% of variance, respectively) but not for friends (11% of variance for both areas of competition). As far as the Social Skills category is concerned, there was a negative, although not significant, correlation with integrative tactics for dating partners (9% of variance); however, for friendships, this area of competition was not correlated with cooperative messages at all. Finally, there was no correlation between the Play category and integrative tactics for romantic relationships. This contrasts with the findings pertaining to same-sex friends, for whom the increase in competing in the area of Play is linked to a decrease in integrative messages, accounting for 20% of variance.

### TABLE 3

<table>
<thead>
<tr>
<th></th>
<th>Distributive Messages</th>
<th>Integrative Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same-Sex Friendships</td>
<td>Dating Relationships</td>
<td>Same-Sex Friendships</td>
</tr>
<tr>
<td>Achievement</td>
<td>.37*</td>
<td>-.23*</td>
</tr>
<tr>
<td>Social Skills</td>
<td>.22*</td>
<td>.00</td>
</tr>
<tr>
<td>Altruism</td>
<td>.37*</td>
<td>-.11</td>
</tr>
<tr>
<td>Social Attractiveness</td>
<td>.40*</td>
<td>-.28*</td>
</tr>
<tr>
<td>Controversy</td>
<td>.32*</td>
<td>-.11</td>
</tr>
<tr>
<td>Affection</td>
<td>.25*</td>
<td>.00</td>
</tr>
<tr>
<td>Play</td>
<td>.20*</td>
<td>-.20*</td>
</tr>
</tbody>
</table>

*p ≤ .01.

Note. Messman and Mikesell (2000) did not include statistical data for the Affection category.

**DISCUSSION**

**Competition and Conflict Interaction Strategies**

The study investigates the link between competition and distributive and integrative conflict messages in same-sex platonic friendships. The categories of competition under scrutiny cover a wide range of behaviors. The respondents assessed the frequency of competing with their same-sex friend over issues such as who is more financially successful, more skillful, talented and knowledgeable. They also reported how often they compete with each other at sports, games and activities as well as during arguments over significant issues. Competing over physical appearance and the attractiveness of social lives were two other aspects rated by the participants.

The results of the current study indicate that individuals’ increased competitiveness in same-sex platonic friendships is related to a greater use of distributive messages and a reduction of integrative tactics. These findings are in line with the aforementioned conflict literature that asserts that competition and interpersonal conflict are intertwined (Cupach & Canary, 2000; Deutsch, 1973; Kohn, 1986; Sillars, 1980).

These results are also consistent with those of Messman and Mikesell (2000), who found that in the case of dating partners, competing in all the areas explored is linked to distributive conflict tactics. Dating couples appear
to compete more frequently in the areas of Controversy and Social Skills, while same-sex friends seem to be more competitive when it comes to Play, Social Attractiveness and Achievement. The Controversy arena, which includes debating, arguing and competing over values, and is most likely to resemble conflict (Messman & Mikesell, 2000), revealed the greatest usage of distributive conflict strategies in dating relationships. This was not the case for same-sex friendships, where Social Attractiveness (having more interesting experiences and nightlife) indicated the highest occurrence of competitive messages. This finding is reflective of the distinct nature of the two relationships and the fact that higher intimacy of romantic couples is likely to lead to more intense emotions and greater conflict (Rawlins, 2001; Voss et al., 1999).

Some differences between the two types of relationships emerged in relation to the link between categories of competition and integrative messages. Namely, for dating partners, competition in the area of Play did not imply a decrease in cooperative messages. This contrasts with same-sex platonic friendships, where competing at sports, games and being funny led to a significant decline of integrative tactics. This is not surprising if one takes into consideration the character of the two relationships and the fact that the time that friends spend together is oftentimes filled with joint activities, such as competitive games and sports (Rawlins, 2001; Voss et al, 1999).

Furthermore, for dating couples, competing in the category of Social Skills decreased the frequency of integrative messages for dating relationships but not for friends. It appears that impressing the other with one’s relational and communicative skills is more salient to romantic partners than to friends. Finally, competition about who does more for the other and who is more giving (Altruism) as well as arguing over important values (Controversy) led to a greater reduction in cooperative conflict behavior in romantic couples than in friends. This mirrors the greater closeness and relational investment within dating relationships as compared to friendships (Voss et al, 1999).

This comparison of same-sex platonic friendships and dating couples does not reveal consequential differences in their competing behaviors or conflict management tactics employed. These findings are not reflective of Emmers-Sommer’s (1999), Laursen’s (1996), and Wright’s (1998) supposition that the nonobligatory nature of friendships and the lack of explicit norms and roles regulating these relationships make them more prone to resort to competitive conflict tactics. Emmers-Sommer as well as Laursen suggest that the higher investment in the relationship and the greater intimacy of romantically involved partners motivate them to utilize a more integrative approach to conflict, which is more conducive to relationship sustenance.

Sex Differences in Competition

The present study also explored the similarities and differences between men and women in their approach to competition and conflict interaction in same-sex platonic friendships. Significant sex differences were found in six out of seven categories of competition examined, with males competing more than females.

Men reported to compete more often than women over issues such as making more money, being more skillful, knowledgeable, successful, receiving more recognition, and performing better at activities (Achievement). They also competed more frequently in being more appealing to the opposite sex than their friend, having more interesting experiences and going to better parties (Social Attractiveness). Arguing and debating over important issues and values (Controversy), as well as competing in sports, games and being funny (Play) were other areas of competition where men rated higher than women. This supports the results of Holmes-Lonergan (2003) and suggests that males indeed get involved in joint activities such as sports and games more often than females and that these pastimes are conducive to their behaving in a dominant and competitive manner. The findings of the present study, which are also in line with Rawlins’ (2001) results, imply that men might more frequently engage in debates with their male friends, make social comparisons such as who is the higher achiever, and be more concerned with status than women.

Women, on the other hand, reported to compete much more frequently than men over each other’s attention (Affection) and issues such as who is a better communicator and who has more relational skills (Social Skills). The inclination of the female same-sex friends to compete over each other’s attention and time might be reflective of their perceiving any interference from others as endangering their relationship (Emmers-Sommer, 1999).

The areas of competition that were relevant to women in the current study corroborate the assertion that females pay more attention to communicative skills and relational maintenance (Rawlins, 2001; Voss, et al, 1999). One of the reasons why they generally compete less than men might be due to the fact that competitive behavior may
Competition and Interpersonal Conflict

be frowned upon by their female friends as being at variance with traditional sex roles (Alagna, 1982). It is also possible that women do not feel comfortable competing with their female friends for fear that it will disturb the relational equilibrium (Benenson et al., 2002; Cheng & Chan, 1999).

In addition, women reported the use of more integrative conflict management tactics than men did, which is consistent with Wright’s (1982) assertion that women’s proclivity to treat their friendships more holistically may lead to a greater regard for the relationship quality and the concern for its maintenance. There were no significant differences between men and women in the use of distributive tactics, indicating that competitive conflict behavior is not restricted to one sex (Meeker, 1990; Richmond & Martin, 1998).

These findings corroborate previously conducted research that asserts the stereotypical dichotomy of competing and conflict resolution behaviors between the sexes (Alagna, 1982; Platow & Shave, 1995; Staley & Cohen, 1988). Both the number and the types of categories of competition where men and women differed are indicative of the traditional distinction that the aforementioned authors found relevant.

Interestingly, in Messman and Mikesell’s (2000) study, there are no differences between men and women engaged in a romantic relationship regarding the frequency of competition and the utilization of integrative and distributive strategies. Their findings substantiate the literature reporting that women are equally competitive and use as many distributive tactics as men (Meeker, 1990; Richmond & Martin, 1998). Messman and Mikesell also discover that women used more conflict management tactics, both integrative and distributive, which supports the argument that they tend to serve as relational experts. They might also be more sensitive to tension and strain that is likely to affect the relationship negatively and are therefore more inclined to make an effort to directly deal with the conflicts that emerge (Wright, 1982).

Although the present study contributes to the understanding of the different nature of two interpersonal relationships: same-sex platonic friendships and dating couples in regard to competitive behavior and conflict messages, caution must be taken to generalize the results to a larger, non-college population. Further research is needed to explore the connection between competition and interpersonal conflict between same-sex friends in other life stages.

The results of the study are further limited by the fact that it focused exclusively on biological sex. Future research investigating competitive behaviors and conflict resolution modes of individuals in same-sex platonic friendships should also take into consideration the psychological gender of participants.

One of the limitations of this study lies in the fact that self-reports were used to evaluate individuals’ behavior during conflict and their proclivity to compete with others. This may lead to subject bias, such as social desirability (Black, 2000). These reports may be inaccurate since people might attempt to present themselves in a favorable light. Thus, further research should address this issue by utilizing both self and friends’ reports as well as observation, in-depth interviews and focus group discussions.

This research could be enhanced by exploring other areas of competing, and making it more explicit to the participants that competition can also include mental comparisons, not just overt behaviors. More information about the friendships might provide further insights into individuals’ approach towards competition and conflict. It would be interesting to explore whether the nature of friendships, their length, intimacy level, and context affect competitiveness or any of its categories.

The study focused on the mainstream American culture, where competition is a salient cultural value (Aronson, 1999; Hirschman, 2003; Tjosvold et al, 2003). It would be noteworthy to study competition in interpersonal relationships in collectivistic societies, where interdependence and the maintenance of harmonious relationships are of paramount importance.

Acknowledgment:

The author wishes to thank Dr. Mark Orbe and Dr. Kathleen Propp for their nvaluabe comments and encouragement.
References


Competition and Interpersonal Conflict


Evaluation of Dye and Pigment Based Ink Jet Ink Sets

VERONIKA CHOVANCOVA
Department of Paper Engineering, Chemical Engineering and Imaging
Western Michigan University

Abstract

The development of high performance inkjet printers and inks is advancing rapidly. Manufacturers seem to introduce their new technology inks to the market on an almost daily basis. Chemists in ink laboratories are still fighting with the issue of combining a wide gamut of dye-based inks with the lightfast and weather resistance qualities of pigment-based inks into new-age ink formulations.

Three different inkjet printers and inks were investigated in this work: the Epson Stylus® Pro 5000, using a dye-based ink set, the Epson Stylus® Pro 5500, employing Archival ink technology, and the Epson Stylus® Photo 2200, with 7-color UltraChrome™ inks. A number of different commercial substrates were sampled. Printability tests were carried out to test and evaluate ink/printer/substrate interactions. Particle size analyses of the three ink types were investigated. Color gamuts and ICC profiles for each of the different printer/ink/substrate sets were compared. In addition, the accuracy of each printer’s color profile was investigated. The results of the profile accuracy measurements were expressed in terms of CIE L*a*b* coordinates and Root Mean Square (RMS) E. Results of accelerated lightfastness tests for the different ink sets were interpreted in terms of change of profile and color gamut.

Introduction

There has been and will continue to be wide development of novel technologies in manufacturing inks and substrates, and due to that, an expansion of inkjet printing technology into desktop, outdoor and industrial applications. Epson has recently introduced two types of pigment-based inks. They combine the advantages of both dye and pigment based inks in their formulations. Both their Archival and UltraChrome™ ink systems represent new ink solutions, where each pigment particle is encapsulated in a resin. This technology offers many advantages over conventional pigment and dye based inks. The primary advantages being those of uniform particle shape and particle size, greater color gamut, advanced optical density, exceptional gloss for photo prints, enhanced lightfastness and support for a wider range of media.

Pigment based inks tend to satisfy the requirements of most ink jet printing demands, but the suitable combination of ink and substrate is still crucial. Inkjet inks require a fine particle size, due to possible clogging of the printing head. For low viscosity inks there is a tendency of particle migration with time. Pigment based inks behave differently than dye-based inks. The spreading behavior of these inks is determined by the hydrodynamic properties such as the Weber or Reynolds’s number. On the other hand, in pigment-based inks, after initial spreading, the pigment particles coagulate on the surface of the microporous layer, creating a filter cake that limits the penetration of the carrier liquid. This results in longer absorption times and recessed dots that stay on the top of the substrate layer, and affect all the other printability properties.

Also, the precision of color reproduction depends on the image processing, e.g. color separation, rendering intent, and on the stability of the printing process, which usually is carried out with the help of an ICC profile and Color Management Modules. In order to understand the whole process, the influence of paper properties on color reproduction has to be taken into consideration. The grade or type of the substrate used will definitely affect the results of the profile calculations and therefore the printing gamut.

Procedures and Results

All the printers (Epson Stylus Photo 2200, Epson Stylus PRO 5000, Epson Stylus PRO 5500) were profiled as CMYK devices on the six selected substrates (Epson Archival Matte, Epson Premium Luster Photo, Epson Premium...
Kodak Glossy, Kodak Satin Paper), using a GretagMacbeth SpectroScanT spectrophotometer (in reflection mode), Gretag-Macbeth ProfileMaker 4.1.5 and the ECI2002 Random Layout CMYK Target\textsuperscript{10}

Sample test prints were produced from Adobe InDesign. In “Color Settings” the CMYK working space was set to the appropriate ICC profile. The prints were made with color management set to source space as proof and the applicable CMYK profile for the print, with the intent set to Absolute Colorimetric for the sample output (the “proof space” is the only management that allows the intent to be manually set). Therefore, all output was set for an absolute colorimetric intent.

**Density Tests**

The samples for all substrates were measured with an XRite 530 SpectroDensitometer. Paper density, Solid density and Dot Area were measured for each sample. The dot area as measured and calculated by the device includes both mechanical and optical gain. Also listed in the results is the difference “Dot Gain” assuming the actual dot size to be a true 20%.

**Particle Size Measurements**

A NICOMP 370 Submicron Particle Sizer was used to measure the particle size of all the ink sets. As expected, no particles were detected in the dye-based ink set for the Stylus PRO 5000. The measured particle sizes of all pigmented inks are found in Table 1.

<table>
<thead>
<tr>
<th>Particle Size</th>
<th>C (nm)</th>
<th>M (nm)</th>
<th>Y (nm)</th>
<th>K (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO 2200</td>
<td>119</td>
<td>172</td>
<td>74</td>
<td>99</td>
</tr>
<tr>
<td>PRO 5500</td>
<td>141</td>
<td>190</td>
<td>123</td>
<td>113</td>
</tr>
<tr>
<td>PRO 5000</td>
<td>Dye</td>
<td>Dye</td>
<td>Dye</td>
<td>Dye</td>
</tr>
</tbody>
</table>

**ICC Profile Test**

Profile accuracy tests were carried out using the following steps. The values of the ColorChecker\textsuperscript{®} target in Photoshop with the profile applied for each paper sample were checked first. This was accomplished by selecting a large portion of each patch and then recording each of the L’a’b’ values from the “Histogram” portion of the “Info” palette. The Mean values obtained from the histogram were converted to actual L’a’b’ values. Using the GretagMacbeth SpectroScanT, L’a’b’ measurements were made for each of the sample patches of the ColorChecker\textsuperscript{®} target for all of the substrates and for each of the sample printers. Employing the formula for color difference \( E^* \),

\[
\Delta E = \sqrt{(L_1^*-L_2^*)^2 + (a_1^*-a_2^*)^2 + (b_1^*-b_2^*)^2} \quad (1)
\]

The original L’a’b’ values of the ColorChecker\textsuperscript{®} target (Target values) were compared with the values from Photoshop with the profile applied (Profile values). These values were also compared with the actual values measured from the printed ColorChecker\textsuperscript{®} portion of the verification samples produced from InDesign, and finally the original values were compared with the values measured from the ColorChecker\textsuperscript{®} Target (Test values). The resultant values for Delta E are listed in Table 2.

<table>
<thead>
<tr>
<th>EPSON Paper</th>
<th>Target vs. Profile</th>
<th>Profile vs. Test</th>
<th>Target vs. Test</th>
<th>IT8/7 Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Photo 2200</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archival Matte</td>
<td>2.42</td>
<td>2.11</td>
<td>2.54</td>
<td>7.55</td>
</tr>
<tr>
<td>Luster Photo</td>
<td>1.48</td>
<td>2.8</td>
<td>2.87</td>
<td>4.39</td>
</tr>
<tr>
<td>Glossy Photo</td>
<td>1.33</td>
<td>1.65</td>
<td>2.02</td>
<td>3.79</td>
</tr>
</tbody>
</table>
The subset part of the IT8/7-3 chart was included in the verification page layout. The L’a’b’ values of the patches were measured with the GretagMacbeth SpectroScanT and compared with the original data of IT8/7-3 chart in order to investigate the quality of the profiles made for each scanner/printer/paper set. The resulting RMS E’s are also shown in Table 2.

**Color Gamut Comparison**

Using CHROMiX ColorThink 2.1.2, the profile gamuts for each of the printers were graphically compared in this order: Epson Photo 2200, Epson Stylus PRO 5000, Epson Stylus PRO 5500 (Figures 1-2). The axis represents the CIELab color space: from “-a” (green) to “+a” (red) and from “-b” (blue) to “+b” (yellow) colors.

![Gamut projection plots for Epson papers, Matte (red), Luster (green) and Glossy (blue) from different printers 2200 (left), PRO 5000 (middle), PRO 5500 (right).](image-url)
Then we compared the similar substrates, glossy and matte/Satin, from each printer to each other. The results were combined and are shown on the 3D gamut plots (Figures 3-4).

**Fading Tests**

The patches of the ECI 2002 Random Layout CMYK Target were measured with the GretagMacbeth SpectroScanT before they were put into the fade meter. They were submitted to 129,600 kJ/m² of energy over 48 hours with the uncoated quartz glass filter configuration and measured again. This represents about 4.5 months (June) of daylight exposure in Florida (36 hrs @ 765 W/m²).

The L’a’b’ values of the printed patches for all the printers on Archival Matte substrate before and after the tests were taken from the data file and the △E calculation was performed to obtain the range of color difference between them (Table 3).
Table 3. $\Delta E$ values before and after fading test for different printers and papers.

<table>
<thead>
<tr>
<th>Printer</th>
<th>Paper</th>
<th>Average $\Delta E$</th>
<th>RMS $\Delta E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photo 2200</td>
<td>Archival Matte</td>
<td>2.20</td>
<td>2.74</td>
</tr>
<tr>
<td>PRO 5000</td>
<td>Archival Matte</td>
<td>10.62</td>
<td>11.34</td>
</tr>
<tr>
<td>PRO 5500</td>
<td>Archival Matte</td>
<td>2.19</td>
<td>2.76</td>
</tr>
</tbody>
</table>

Table 3 does show that the pigmented inks change colors much less than the dye inks, as expected. However, values $\sim 3$ for the pigmented inks are larger than expected for inks rated at more than 75 years$^{13,14}$. Examination of the data shows that there is a systematic shift toward yellow and green. The Epson 2200 shows an average $b^*$ of 1.57, while the Epson 5500 shows an average $b^*$ of 1.89. Thus, for the pigmented inks, most of the average $\Delta E$ results from the systematic $b^*$ shift, reflecting the drop in the OBA$^{15,16}$ contribution (see below). The Epson 5000 shows an average $b^*$ of only .77, but the average $L^*$ is 6.96. Therefore, that $\Delta E$ is mostly due to actual ink fading.

Again, the profile gamut plots for the papers are given in Figure 5. Figure 5 shows the gamut plots before and after the fading test.

Fig. 5. Comparisons of projections of the color gamuts before (full color) and after (black) fading test for pigment-based Epson 2200 (left), dye-based Epson 5000 (middle) and pigment-based Epson 5500 (right).

Note that the Epson 5000 shows a significant decrease in color gamut because of the dye-based nature of the used inks. The printers with the pigmented inks, the Epson 2200 and 5000, show the aforementioned shift towards yellow, but little decrease in gamut.

The Epson Stylus Photo 2200 printer together with the Epson Archival Matte substrate provides best results for this part of the research. This set was chosen for further investigation of the fading properties. This substrate with the printed chart from the 2200 was submitted to longer time light exposure equivalent to 13 months (June) of daylight exposure in Florida (104 hrs @ 765 W/m²). The gamut plot of this test is shown in figure below. In this case, the color shift is even more significant in the yellow region of the spectrum.
From the gathered information we decided to look at the changes in properties of the plain substrate. \( L^*a^*b^* \) values of the substrates before and after the tests were taken. \( \Delta E \) calculations for obtaining the range of color difference are shown in the Table 4.

Table 4. Average and RMS \( \Delta E \) values before and after fading test for different printers and papers.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>L*</th>
<th>a*</th>
<th>b*</th>
<th>( \Delta E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epson Archival Matte</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>96.1</td>
<td>0.8</td>
<td>-4.3</td>
<td>4.34</td>
</tr>
<tr>
<td>After</td>
<td>95.8</td>
<td>-0.4</td>
<td>-0.1</td>
<td></td>
</tr>
<tr>
<td>Kodak Satin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>93.3</td>
<td>0.7</td>
<td>-6.3</td>
<td>2.49</td>
</tr>
<tr>
<td>After</td>
<td>93.4</td>
<td>-0.1</td>
<td>-3.9</td>
<td></td>
</tr>
<tr>
<td>Epson Premium Glossy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>94.6</td>
<td>-0.4</td>
<td>-3.9</td>
<td>0.50</td>
</tr>
<tr>
<td>After</td>
<td>94.4</td>
<td>-0.6</td>
<td>-3.5</td>
<td></td>
</tr>
<tr>
<td>Kodak Glossy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>92.8</td>
<td>0.3</td>
<td>-6.7</td>
<td>2.66</td>
</tr>
<tr>
<td>After</td>
<td>93.7</td>
<td>0.1</td>
<td>-4.2</td>
<td></td>
</tr>
<tr>
<td>Epson Archival Matte</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(long term test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>95.9</td>
<td>0.8</td>
<td>-4.0</td>
<td>4.91</td>
</tr>
<tr>
<td>After</td>
<td>95.8</td>
<td>-0.6</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

The GretagMacbeth MeasureTool 5.0.0 software was used to compare the spectra of the substrates before and after the fading test. The spectra for the Epson Archival Matte substrate, claiming the best archival properties, Epson Glossy substrate, Kodak Glossy substrate and for Kodak Satin substrate are shown in Figures 7 to 10.
The spectra and the L’\(a^*\)\(b^*\) values suggest that the contribution of optical brighteners, added to improve the perceived whiteness of the paper, has been neutralized for the Archival Matte paper and greatly diminished for the Kodak Satin papers. Optical Brightening Agents (OBA) are fluorescent materials that absorb in the ultraviolet and emit in the blue\(^{15,16}\). This is the source for the blue peak in the spectra and the negative values of \(b^*\) before the fading test. This means that, regardless of the permanence of the printed dye or pigmented ink, there will always be some shift in the perceive color of printed images. Note from Table 4 that the majority of the OBA neutralization has occurred in the first simulated 4.5 month period, with little (barely significant) additional change in the remaining simulated 8.5 months.

**Other Properties of Printer/Substrate Combinations**

Other properties of the Printer and substrate combinations are given in other research paper. In particular, the paper roughness by Parker Print Surf, profilometer and Atomic Force Microscopy\(^{17,18,19}\).

**Discussion**

The procedures used with the densitometer and the measurements obtained by that method produced comparative values for the 20% dot area on all of the samples, and all measurements were of comparative values for all of the papers and inks. The matte samples from all printers did represent a lower density than those of the luster and glossy samples. The dot gain seemed relatively consistent for all colors on all samples.

The comparison of the difference in \(\Delta E\) values for the original L’\(a^*\)\(b^*\) ColorChecker\(^{®}\) target to those of the values calculated in Photoshop indicate small dissimilarities in almost all cases. The \(\Delta E\) values for most of the patches on all substrates and from all printers were found to be generally less than two. Exceptions include the dark patches when printed on the matte papers and when printed from the Photo 2200 and PRO 5500 using pigment based inks. In the case of the pigment based ink printers (Epson 2200 and Epson 5500) the average and RMS \(\Delta E\) were always higher for the matte substrates than for the luster, satin and glossy substrates. This is most likely due to out of gamut colors for the matte substrates.

The \(\Delta E\) values for the comparison of the patches calculated in Photoshop to those measured with the SpectroScanT show similar values to the differences between the original values and the values from Photoshop in the case of the Epson papers. The only exception is the Epson Stylus PRO 5000 in combination with Kodak substrates.

Comparisons of the measured samples in most cases very closely approximate the values of the original ColorChecker\(^{®}\) reference values, with the largest variances indicated on the glossy papers printed from the PRO 5000 and the matte from the PRO 5500. Matte paper printed from the PRO 5500 produced the largest variances of all the samples.

In comparing the profile gamuts it was noted in all cases that the matte paper profile represented the smallest gamut whereas the luster and glossy papers were generally similar and contained the complete matte gamut. Comparing the printers to each other on the same substrate the Photo 2200 generally included a similar size gamut to that of the PRO 5000 printer and dye based inks but the PRO 5500 represented the smallest color gamut. It could be seen that the Photo 2200 with its pigment-based inks is able to provide a color range that very closely matches that of the dye based prints from the PRO 5000.

The smaller gamut produced by the PRO 5500 printer may have something to do with the older technology and/or the advertised better archival properties of the ink set used by that printer. The fact that the pigment based
inks used in the Photo 2200 printer closely match those of the dye based inks of the PRO 5000 is noteworthy, but it can be expected that the archival properties as advertised for this ink set may not be as good as those of the PRO 5500. It should also be noted that the increased archival properties of the matte paper in combination with archival pigment based inks produce the smallest color gamut of the samples analyzed.

Taking into the consideration the Kodak paper, there is no difference in gamut size between glossy and satin substrate. In addition, Epson vs. Kodak paper gamuts did not show any significant discrepancies in the terms of color gamut size. It is seen from Figure 3 and 4 that the widest gamut was obtained when printed from the Epson Stylus PRO 5000 dye based inkjet printer followed by Epson Photo 2200 and Epson Stylus PRO 5500, both pigment based inkjet printers.

After the printouts were submitted to the fading test it could be seen that the gamuts decreased. The Epson 5000 showed a significant decrease, while the 2200 and 5500 showed small changes. In the case of the Epson Archival Matte and as well the Kodak Satin substrate, it was found that, even without any change in ink composition, the color performance will change because of the loss of brightener effect. This led to a systematic shift toward the yellow, especially when exposed to longer time tests, as shown in Figure 6. This deviation was not seen when inspecting the glossy substrates.

The particle size of the pigment based inks were found to be in the range <190 nm, most of them bellow 150 nm, showing smaller particle sizes for the PRO 2200 ink set than for the Photo 5500 ink set. The Particle Sizer’s light detector was not able to distinguish any intensity in the case of the PRO 5000-ink set, which is consistent with the dye based ink system of the printer. The color gamut decreases with particle size, with the smallest particle size, the Epson 5000 dye, having the largest gamut, while the largest particle size, the Epson 5500, gives the smallest gamut. However, the dye based ink in the 5000 showed significant fading from only a simulated 4.5 month exposure.

**Conclusion**

Different inkjet printers and their corresponding ink sets were studied in terms of printability tests, ink/printer/substrate interactions, particle size analyses, color gamut comparisons, the accuracy of printer’s color profile, and fading tests. It can be definitely said that the new technology of the manufacturing the inks with pigment particles encapsulated in specific resins is able to approach the properties of the dye based inks, especially in the term of gamut width. The particle size of the pigment in these inks is small enough to provide the color range that could match that of the dye based inks and also reach the gamut of digital silver halide photo on conventional photo paper. Also, it has to be mentioned that the increased archival properties of the matte paper in combination with archival pigment based inks reflect in the smaller color gamut than the gamut of glossy paper. The pigment based inks show much better lightfastness than the dye-based inks, but for some substrates there is a drift towards the yellow as optical brighteners lose their effect.

For future work we suggest to investigate the substrates which do not include optical brighteners in their composition, e.g. art paper. Also, there is a newer dye based ink set becoming available, with enhanced archival properties. HP has created a new generation of inks to achieve over 100 year predicted indoor lightfastness performance, while simultaneously improving the color gamut over previous products.
References

In order for a naturalized account of epistemology to be taken seriously by the philosophic community, it must address questions raised by the skeptic. Since W.V. Quine’s essay in 1969 entitled “Epistemology Naturalized,” the naturalized project has been marked both by bold claims that the traditional project has failed and that the problems a naturalized project potentially faces are not problems of concern. Instead of addressing these potential problems head on, attempts to advance the naturalized project since Quine have largely masked their language, consequently disguising the boldness of their separation. In an essay entitled “Against Naturalized Epistemology,” Laurence BonJour brings to light the shortcomings of the naturalized project while heavily criticizing Quine. Further, BonJour criticizes the contemporary naturalist Philip Kitcher. BonJour persuasively argues that Quine has failed to offer cogent reasons for abandoning the traditional project in favor of a naturalized project, while maintaining that Quine has also failed to offer a viable alternative. Further, BonJour argues that Kitcher’s reasons for psychologizing epistemology are innocuous. Finally, BonJour outlines an argument that concludes that the abandonment of a priori justification leads to epistemological “disaster.” While BonJour’s arguments are at times convincing, they are not wholly successful. It shall be the purpose of this paper to refute the crux of BonJour’s arguments, while at the same time attempting to resolve the problems he claims a naturalized project faces. It is only after these problems have been addressed directly that a complete naturalized epistemology can be developed.

The distinction between internalism and externalism is a simple but decisive one, which hinges on what epistemic principle can serve as the appropriate connection to truth. For the internalist, all of the epistemic weight lays with rationality. Conversely, the externalist claims that rational justification does not constitute knowledge, and therefore, success ought to replace it as an epistemic principle. To make the distinction clear, consider the following scenario. Suppose, for instance, that aliens were to beam information about the outcome of football games into a man’s head. Come Sunday morning, and without knowing exactly how or why, he suddenly has a host of beliefs about which teams will win every game. Several weeks pass, and without exception, his predictions are correct. Can his beliefs about which teams will win next Sunday be considered knowledge? For the internalist, the answer is an obvious no. Again, the internalist is concerned with rationality and thus demands reasons for the belief at hand. Externalists have, however, failed to offer a compelling argument in favor of success as an epistemic principle. Rather, they have only provided a scenario in which success serves as the epistemic principle. On the other hand, the internalists hold that no empirical claim can serve as an epistemic principle, in that claims of reliability are contingent. For the internalist, epistemic principles are a priori, and contingent claims are obviously not a priori.

This is precisely the point that BonJour makes. He writes, “What naturalism needs, then, is a direct objection to rationalism—or, more or less equivalently, an argument that the idea of a priori justification is untenable, even when not construed in a moderate empiricist way” (BonJour). Further, the externalist has failed to consider the ramifications of arguing against a priori knowledge, as any appeal to such knowledge would be a concession to internalism. Specifically, BonJour claims that it is the renouncement of a priori knowledge that renders the concept of naturalism a disaster. And so, if the externalist is to succeed s/he must not only offer evidence that shows that his stance has the theoretical capacity to answer questions raised by the skeptic, but s/he must also show that the position taken by the traditional epistemologist is an unfruitful one. Further, the problem of circularity that emerges as a result of taking an external position must be resolved. The problem of circularity was recognized by Quine in 1969; however, it is widely held that a satisfactory response has not yet been offered.
Internalism, though, is not without problems of its own. Specifically, internalism faces Hume’s problem of induction as well as all Cartesian type problems. While it has been argued that these problems are fatal to the traditional project, the internalist has offered adequate responses. To Cartesian problems such as illusion and evil deceivers, the internalist is not invulnerable. However, certainty is not a requirement for knowledge. Rather, only rationality is required. Internalists offer no guarantee but merely elucidate what is logically possible. If agents are unknowingly deceived, it is through no fault of their own that the world has failed to cooperate. In regard to Hume’s problem of induction, it seems apparent that Hume has misconstrued the requirements for knowledge. Further, today’s epistemologist has a variety of tools at hand that were unknown to Hume. Using incremental probabilities and Bayesian Theorem, Dr. Timothy McGrew has offered a compelling response to the problem of induction in his essay “Direct Inference and the Problem of Induction.” And so it seems that the externalist is left with two primary options: (i) knowingly offer what can only be an alternative theory of knowledge to the traditional project while ignoring all of the standards it has established, or (ii) provide a new argument against a rational connection to truth, rendering it untenable. Option (i) has, since Quine, been the most common tactic employed by the naturalist. Option (ii), on the other hand, can only be pursued via an attack on a priori knowledge.

BonJour, however, delineates an argument against externalism and in favor of a priori justification that, if true, would render the naturalistic project a “disaster.” BonJour writes:

If we are to have any reason for thinking these latter beliefs to be true, such a reason must apparently either (i) depend on an inference of some sort from some of the directly observational beliefs or (ii) be entirely independent of direct observation. A reason of sort (ii) is plainly a priori. And a reason of sort (i) can only be cogent if its corresponding conditional, a conditional statement having the conjunction of the directly observational premises as antecedent and the proposition that is the content of the non-observational belief as the consequent, is something that we in turn have a reason to think to be true. But the reason for thinking that this latter, conditional statement is true can again only be a priori: if, as we may assume, all relevant observations are already included in the antecedent, they can offer no support to the claim that if that antecedent is true, then something further is true (BonJour 11).

In other words, if one is justified in believing premise one, premise two, and all the way down to premise x, in that they are known empirically, then one’s justification for believing the conclusion (not known through observation) can only be a priori. BonJour claims this to be an epistemic principle. This certainly seems to be a damaging argument against the naturalist, as all of the empirical information is packed into the antecedent. Thus, no way exists to reach the conclusion without an appeal to knowledge known independent of experience.

Though BonJour concedes that empirical evidence may indeed do most of the epistemic work, he maintains that a priori knowledge serves as the link between a collection of observations (the antecedent) and the conclusion (the consequent). But, is it obvious that such a link must be a priori? Further, how does one come to have a priori knowledge? A priori knowledge in the form of an inference is used when one’s knowledge is incomplete, such that a conclusion not observed in experience can be reached. This is synthetic a priori knowledge, which BonJour argues in favor of, as opposed to analytic a priori knowledge in the form of tautologies, which are true of all possible worlds. The latter seems undeniable, though trivial, while the former is precisely what is in contention. That inferences can be demonstrated to be likely to a degree via probability also seems undeniable. We are able to delineate a set of logically possible conclusions, as well as to give each possible conclusion a probability of likelihood. Whether this sort of method should count as knowledge is the heart of the question. While it is common for naturalists to argue that a probability cannot serve as knowledge, the objection this paper makes runs much deeper. The objection is that this sort of demonstration amounts to nothing more than reflection upon experience. It is not that it should be disregarded. Further, BonJour’s assertions allow for the possibilities of such an objection.

In responding to an argument advanced by Philip Kitcher, BonJour offers an alternative conclusion to the possible conclusions outlined by Kitcher. BonJour’s alternative, though, seems to put him in a precarious position. In the following passage, BonJour quotes Kitcher in response to a Kuhn-inspired argument:
Either they can continue to insist that philosophers know a priori the principles of confirmation and evidence, concluding that the actual reasoning of scientists is cognitively deficient, or they can abandon the a priori status of methodological claims and use the performances of past and present scientists as a guide to formulating a fallible theory of confirmation and evidence. Since the first option has an uncomfortable air of arrogance, it is hardly surprising that most responses to Kuhn have followed the latter course (BonJour 9-10).

In response, BonJour claims that the first two alternatives: “(i) sticking stubbornly and perhaps dogmatically to one’s initial allegedly a priori precepts” and “(ii) adopting what may seem to amount to an abandonment of philosophy altogether in favor of a kind of psychology or sociology of scientific practice” are not exhaustive. As a third option, BonJour purposes that:

…reconsidering one’s initial a priori assessment in light of what scientists seem to be doing, while still insisting that any acceptable mode of scientific reasoning must ultimately be able to be seen or shown to be rationally cogent in an a priori way that transcends the mere fact that it is employed in practice (BonJour 10).

But, can this third option still be construed as a priori? Is it not based on observation? This seems to beg the question. If a priori knowledge is defined as being independent of experience, then BonJour’s sort of a priori knowledge seems to be dependent upon experience and thus fails to be a priori. On one hand, he claims that a priori knowledge is necessary to reach a consequent from an antecedent packed with observation, and on the other, he concedes that a priori knowledge must be shaped by what is observed. This seems to be an obvious contradiction. In fact, BonJour admits as much, even going so far as to refer to a priori knowledge as both “fallible and corrigible.”

BonJour writes:

If this third alternative is indeed available, then Kuhn’s historical argument fails to establish the strong thesis that the ideas of a priori justification and a priori epistemology should be abandoned, as opposed to the much weaker thesis that an apparent conflict with scientific practice can and should provoke a reconsideration of the specific a priori results in question, possibly but certainly not inevitably leading to their correction or abandonment. All this assumes, of course, that it makes sense to speak of correcting or abandoning a priori claims, i.e., that a priori justification need not be construed as incorrigible or infallible (BonJour 10).

To hold a priori knowledge as both fallible and correctable seems equivalent to labeling it a priori as a necessity and in order to maintain a philosophical view, rather than because it is indeed a priori. However, if a priori knowledge is in fact dogmatic, then Kuhn’s objection remains and BonJour has failed to offer a viable alternative.

For the externalist, however, success can serve as an adequate connection to truth. But, until a sufficient track record of trial and error has been established, the externalist is left only with a true belief. With a proven history of success in hand though, the externalist is armed with justification on an objective level. It is just this sort of trial and error process by which BonJour seemingly admits that “a priori” knowledge is manipulated around changing scientific methods such that a sufficient track record is established. In other words, BonJour seems to want a priori knowledge to be correctable in light of mistakes and in accordance with scientific findings. If it works, then use it, but refer to it as a priori. Here, the traditional epistemologist has fuzzy word definitions and may even be guilty of changing the subject in order to achieve desirable results. This clearly demonstrates that what has traditionally been referred to as a priori knowledge is nothing more than a reflection upon experience. This is defined as concept empiricism. BonJour is guilty of using a loose and flexible definition of a priori that can be stretched to yield the desired results.

If a priori knowledge as an epistemic principle is untenable in any meaningful way, then the externalist is in need of a viable alternative. The obvious alternative to a priori knowledge is empiricism. Arguing in favor of empiricism is certainly nothing new and would put one in the company of philosophical giants such as Locke and Hume. However, there are numerous forms of empiricism, each with problems. First, though, one must have an adequate understanding of the concepts at hand, as well as an agreed upon set of working definitions so as to not be accused of changing the subject. Victory by way of definition would be a hallow one, and this seems to be the
only progress made by naturalism thus far. A priori knowledge is commonly defined as being independent of any justification from experience. However, experience from which the concept at hand is derived may be presupposed. For example, most philosophers consider mathematical knowledge to be a priori, though children cannot acquire such knowledge until experience of the world is first gained. Knowledge, however, which can only be justified by at least some appeal to experience, i.e. the five senses, is referred to as empirical knowledge (a posteriori). BonJour is a judgment rationalist, claiming that there are synthetic truths knowable a priori. Concept empiricism, on the other hand, states that every concept is either given in experience or is generated by reflection on experience. This is the position held by Locke and Hume, and in light of BonJour’s concessions, it seems to leave the door open for the naturalist to hijack what has historically been referred to as synthetic a priori knowledge. If it can be shown that knowledge historically considered to be a priori is generated by reflection on experience rather than being known independently of it, then the internalist’s position is weakened, and room is left for a form of externalism.

In his essay “An Enquiry Concerning Human Understanding,” David Hume undermines the notion of a priori knowledge. Hume writes:

In a word, then, every effect is a distinct even from its cause. It could not, therefore, be discovered in the cause, and the first invention or conception of it, a priori, must be entirely arbitrary. And even after it is suggested, the conjunction of it with the cause must appear equally arbitrary; since there are always many other effects, which, to reason, must seem fully as consistent and natural. In vain, therefore, should we pretend to determine any single event, or infer any cause or effect, without the assistance of observation and experience… (Hume 436).

Concept empiricism, however, is not without problems of its own. Hume conceded that this view was not entirely correct, as was demonstrated by the missing shade of blue. Imagine a spectrum of various shades of blue such that each patch of blue in the spectrum gradually and progressively becomes darker. Is it possible to imagine a missing link in the spectrum even though that exact shade is not presented to us in experience? The answer seems to be an obvious yes, and consequently concept empiricism needs to be revised. However, the missing shade of blue, though a concept not given to us in experience, is derived from experience. Though it seems obvious that it is indeed possible to imagine the missing link, the imagination provides only the missing link and not the spectrum in its entirety. What the would-be naturalist is in need of is an argument that clearly demonstrates that what has traditionally been referred to as a priori knowledge is generated by reflection upon experience. Adopting such a position would allow for the naturalist to incorporate what is generally held to be a priori knowledge by arguing that a priori knowledge is not independent of experience after all but is, rather, dependent upon it and is thus, by definition, not a priori.

However, given BonJour’s concession that a priori knowledge is fallible, is it conceptually plausible for the role that a priori knowledge has traditionally played to be replaced by a form of emergentism? Further, what are the consequences of abandoning a priori knowledge? If analytic a priori knowledge is trivial, and even if synthetic a priori knowledge is not truly a priori, then it still does not seem as though the externalist is equipped with the necessary tools to solve BonJour’s problem and bridge the gap between antecedent and consequent.

There is, however, a third type of “a priori” knowledge in addition to analytic and synthetic knowledge. Kant talks of a priori knowledge as knowledge that cannot be taken from experience in that, without such knowledge, we would not be able to make sense of the world. He asks what it is possible for any human-like mind to know, concluding that we must first critique our “faculties.” This is perhaps the only way in which it makes sense to speak of knowledge as being truly a priori. Clearly, a mind void of certain abilities, i.e. the ability to pattern match, would not be able to function in the world. While it is well beyond the scope of this paper to enumerate such capabilities, their existence seems undeniable. Though Kant’s classification of a prior knowledge would appear favorable to the internalist in that it is a form of knowledge independent of experience, I believe that it has no bearing on the traditional project whatsoever. It says nothing about reasons for holding a belief but, rather, elucidates how it is possible for beliefs to be formed. And so, a priori knowledge in a Kantian sense can only strengthen the externalist’s position.
BonJour’s critique of naturalized epistemology is that, given a set of empirical premises packed into the antecedent, there is no way that is not a priori to reach a consequent observed in experience. But, it seems that Kant’s definition of a priori knowledge has left open a possibility for the externalist to argue his way out the back-door. While many naturalists have attempted to argue that the ability to reach true beliefs a high percentage of the time has been hard-wired into the brain in an attempt to circumnavigate the problem of circularity, all such attempts have failed thus far. However, given a Kantian view of a priori knowledge, no such appeal is necessary. In fact, it seems plausible that given a small number of a priori “categories,” to again borrow Kant’s terminology, that a logically ordered world-view will begin to emerge given sufficient empirical input. This does not directly refute our previous definition of a priori knowledge, which hardly seems to be controversial. Even mathematical knowledge, widely held to be a priori, is dependent upon experience of the world. Further, it is not at odds with BonJour’s admission that a priori knowledge is fallible. In fact, it insists upon it. What is significant, however, is that given such a view, a priori knowledge is by and large shaped by experience, with the only exception being a priori knowledge in the Kantian sense. Further, such knowledge gives the externalist an a priori foundation by which to bridge the gap between antecedent and consequent. It is also consistent with concept empiricism.

The need to construct a consistent world in which cause and effect relationships hold true seems to be a psychological necessity of survival. Try to imagine a world in which such relationships do not exist. Information received via the five senses must be organized in a logical manner in order to make sense of the world. Even if the cause and effect relationships that we project onto the world are not an accurate representation of it, they seem to function within it. And so, if the ability to pattern match, etc. paired with memory yields a consistent, logical domain as empirical knowledge of the world is gained, then the emergence of the missing shade of blue as a reflection upon experience does not seem to be particularly mysterious when given the rest of the spectrum. Determining what specifically constitutes Kantian a priori knowledge, however, is a job for cognitive science, psychology, artificial intelligence, and perhaps transcendental arguments.

However, another objection remains. While the internalist many not particularly disagree with this view, he would, I believe, object on the grounds that none of it addresses the questions raised by the skeptic. Thus, it fails to count as epistemology and is therefore in no way philosophically interesting. Rather, the only accomplishment, if anything at all has indeed been accomplished, is an understanding of how beliefs are formed. Remember, the argument at hand is one of metaepistemology. That is, what are the correct standards for justification in a Knowledge = True Justified Belief format? The internalist is entitled to only what an agent is directly aware of and can “know” from within, while the externalist throws consciousness out the window and is left only with stimulus and response. Success, however, is not necessarily a result of being rational, and being rational does not guarantee success, objects the traditionalist. However, if the only thing that separates the internalist from the externalist is a priori knowledge as an epistemic principle, and this epistemic principle has not only been shown to be fallible, but is dependent upon experience to even be formed, then the traditional project must by definition collapse into the naturalized project. In regard to skeptical questions, the externalist is no better off. Certainty remains impossible. However, success can be demonstrated and can knowingly be achieved in that we are able to survive and navigate in a matrix of uncertainty.

A final problem arises for the naturalist as a result of an external account of justification. BonJour writes:

…one thing that it is important to bear in mind about the issue of a priori justification is how easy it is to rely on a priori insights without explicitly acknowledging them, even to oneself. This is particularly easy where such insights to fundamental patterns of reasoning and argument (BonJour 13).

Using success as the epistemic principle, which connects us to truth, generates a problem of circularity. If the goal is to sort true from false beliefs, then we are in need of a reliable process capable of doing the sorting. The question then becomes which method, or process, shall we use. The answer is obviously one that yields true beliefs, and by examining the results of the process, we can determine if they are true. But how do we know that they are indeed true? After all, this was the purpose of the reliable process in the first place. And so it seems that we have entered into a hopeless problem of circularity.
Quine offers a possible, though problematic solution:

Such a surrender of the epistemological burden to psychology is a move that was disallowed in earlier times as circular reasoning. If the epistemologist’s goal is validation of the grounds of empirical science, he defeats his purpose by using psychology or other empirical science in the validation. However, such scruples against circularity have little point once we have stopped dreaming of deducing science from observations (Quine 290).

In other words, Quine argues that the problem of circularity really is no longer a problem, at least not one worth addressing. Ignoring the problem, however, does not make it go away, and until it has been adequately addressed, the traditional epistemologist will have grounds from which to object to a naturalistic project and reasonably so. The problem of circularity is a problem of criterion. How does one deem a process a successful one? How does one know that a process is a reliable one? However, the needed criterion is a simple and obvious one.

The naturalist is in need of an established track record by way of trial and error that confirms the process as a reliable one. This seems to be an accurate representation of what we experience, as well as a matter of survival. My senses must correlate with the external world to a sufficient degree such that I am successfully able to navigate within it. As Allan Gibbard put it in his book Thinking How to Live, “Clearly those who were hopeless at getting everyday, surrounding facts right would tend not to reproduce. They would bump into things; they would fall over cliffs and get killed” (Gibbard 255). While I agree with Gibbard’s claim, it is this sort of reasoning that the traditional epistemologist objects to. It must be noted that the epistemological question of how our claims to knowledge can be justified is different from the psychological question of how we come to hold these beliefs. However, if the epistemological question, as has been defined by the traditionalist, is untenable, then the latter alternative may be all with which we are left.

Unfortunately, this line of reasoning only leads to further objections. First, it is unclear that given this sort of criterion, any interesting results may be achieved. Second, it does not seem to answer questions raised by the skeptic but, rather, answers questions about how beliefs are actually formed. Third, it results in justification across a period of time rather than at a particular moment. The traditional epistemologist is concerned only with the justification of beliefs at a particular time. The problem for the naturalist thus becomes: belief X for person A can be justified by the process by which person A arrived at belief X, while it is possible that the same belief X for person B is not justified in that it was not formed as the result of a reliable process. On this account, a naturalized justification of knowledge yields a nonsensical result. Either belief X is justified for both person A and person B, or it is not justified for both person A and person B. Finally, a host of objections specific to a Darwinian evolutionary process have been raised and must be addressed individually.

However, if a priori knowledge is both fallible and correctable, as BonJour concedes that it is, then by what means does one know when a priori knowledge is infallible? This seems to put the internalist in the same position as the externalist in regards to circularity. If a priori justification is reliant upon empirical evidence to determine its validity and empirical evidence is justified by success via some sort of reliable process, then circularity is problematic for both accounts of justification.

There are a number of levels on which this paper can potentially make progress: (i) by showing that BonJour’s attack on the very possibility of externalism has been refuted, (ii) by showing that the hopes of a naturalized project have not been rendered a disaster, and (iii) by laying the foundation for a complete naturalized project by delineating how it might be possible to refute internalism as well as addressing the problem of circularity. If, however, this sort of attack employed against a priori knowledge is conceptually invalid, then the project of naturalized epistemology does not seem to be valid. But if it is successful, there are still significant problems that remain for a naturalized epistemology to overcome. These problems though are beyond the scope of this paper. Rather, I hope only that a potential line of reasoning for a naturalized project has been established.
Notes

i BonJour defines moderate empiricism as “The thesis that the only claims or propositions that can be justified a priori are those which are analytic” (BonJour 8). Further, BonJour notes that, at most, it is this reduction to analyticity that Quine has argued against successfully. He writes, “…Quine’s arguments against the idea of a priori justification tend to assume what amounts to a hypothetical version of moderate empiricism: the view that if there were any a priori justified claims, they would have to be analytic…It is, I suggest, this moderate empiricist argument, and not the idea of a priori justification itself, that is defeated by Quine’s arguments (assuming that they are otherwise successful)” (BonJour 8).

ii Speaking of a priori epistemic principles, BonJour writes, “Nothing about the argument advanced here excludes the possibility that such principles might be few in number and very general in character, perhaps even limited to logic (including probability theory) and general principles of inductive and explanatory reasoning. Thus it might be that such principles do not take us very far-and in particular that much of the job of ‘meliorative epistemology’ must be done empirically” (BonJour 13).

iii Though much of Kant’s philosophy concerns his defense of synthetic a priori propositions, I wish to not only make a distinction between synthetic and analytic a priori knowledge but to also treat Kant’s conception of “faculties” as a third a priori category, however controversial this distinction may be.

References

Comparison of Computational and Experimental Aerodynamics Results for a WMU Solar Car Model

YANG YANG
AND
WILLIAM LIOU, PH.D.
Department of Mechanical and Aeronautical Engineering
Western Michigan University

ABSTRACT

Three-dimensional numerical simulations using FLUENT [1] were performed to model the airflow over the Sunseeker, an award-winning solar car that was designed and built at Western Michigan University. Converged numerical solutions on three different grids are reported and compared with the available experimental data, which include the lift and the drag coefficients. Also reported are the results obtained by using the second-order upwinding discretization on one of the grids. The comparison shows that the computed lift coefficients agree well with the experimental data for all the three grids and the different orders of numerical methods, indicating that the pressure field is well captured. The agreement with the data for drag coefficient varies, which appears to suggest a higher degree of dependency on the grid distributions than that for the lift coefficient. These results are discussed in terms of their implications for the simulations of similar low-drag vehicles.

INTRODUCTION

Western Michigan University’s (WMU) Sunseeker solar racecars have participated in a series of American Solar Challenge, a biannual competition with about 20 teams participating in each race since 1990. These represented colleges and universities from across the United States and Canada. In the latest race in July 2003, the Western team earned fifth place in the 2,300-mile race.

The WMU Sunseeker has three main parts: the canopy, the upper surface that is covered with an array of solar battery cells, and the chassis. Figure (1) shows a photo of the 2003 team and the Sunseeker car that was simulated here. The Sunseeker is 196.25 inches in length and 68.75 inches in width. The top speed can reach 70 mph running solely on the solar energy provided by the solar panels on the upper surface.

One of the emphases considered when the solar car was designed was achieving low aerodynamic drag. The design also must provide enough upper surface area for the solar panels and enough volume space to house the driver and the motor. Wind-tunnel testing of the Sunseeker model has been performed at two industrial facilities and also in the Applied Aerodynamics Laboratory at WMU. In the WMU subsonic wind tunnel, a 1/5 scale model was used. Figure (2) shows the setup in the test section of the tunnel. The model was bolted down to the floor of the test section with no wheels. A six-component force balance was used, and the data for the forces and the moments were obtained [2]. A smoke gun has been used in some cases to provide a visualization of the flow. There is, however, no detailed measurement on the flow properties, such as the surface pressure and the velocity distributions. There are also no data that indicate whether the boundary layers are laminar, transitional, or turbulent at the air speeds tested.

This paper reports the results of the numerical solutions of the three-dimensional flow field around the Sunseeker. Computational fluid dynamics (CFD) simulations of the flow fields around road vehicles are critical to modern-day automobile designs and many articles have been written in this area. A comprehensive review can be found in Hochu and Sovran [3]. CFD has also been used in the development of the aerodynamics characteristics of solar car [4]. Compared with the traditional automobiles that run on internal combustion engine, the energy available for solar cars is far less. Therefore drag reduction must be a main consideration in the aerodynamic design of solar car.
From the vehicle stability point of view, it is also important to evaluate the amount and the direction of the lift. With the resources available, the WMU Sunseeker has been designed with limited use of CFD tools. Although the car has performed satisfactorily in the past races, it is important that matured CFD technologies be assessed and considered for the future designs.

The widely available commercial software FLUENT is used in this study. The computed aerodynamic characteristics of the Sunseeker, including the lift coefficient, the drag coefficients, and the surface pressure distribution will be reported and compared with the wind tunnel data. In addition, the simulated flow will be studied by examining the flow pressure and velocity distributions around the car.

**SIMULATION CONDITIONS**

The computational model of the Sunseeker was built based upon an IGES file that was generated by a CAD software. The topology and connectivity information about the car surfaces did not affect the manufacturing of the car parts and were not recorded in the CAD file. Therefore, the geometric information retained in the IGES files is not sufficient for CFD. To improve the quality of the geometry, we have used Pro/E to reconstruct the model. The disconnected edges and faces were “stitched” together and some poorly defined surfaces were replaced by mathematically smooth curves. As it is well known, the initial meshing of a new geometry is the most time-consuming part of a CFD simulation, when the model has to be “cleaned up”.

After the geometry was reconstructed and became numerically acceptable, the file was then imported into Gambit, an accompanying meshing software of FLUENT, to construct a three-dimensional grid system. Since this is a symmetric three-dimensional external flow problem, only a half of the Sunseeker needs to be simulated. A rectangular computational domain, as shown in Figure (3), has been used. The domain accommodates the half-car model with a symmetry plane. The car surfaces were first covered with uniformly distributed triangular mesh elements. The surfaces, combined with those from the outer block, were stitched together to form a single meshable volume of the computational domain. Tetrahedral unstructured mesh cells were used in the computational domain.

Three different grids, denoted as Grid A, Grid B, and Grid C, were used in this study. Grids A and C both employ the size function from Gambit, while Grid B uses the stretching scheme. For Grid A, the cell size gradually grow from around 0.0069 m near the car surface up to about 0.271 m in the far-field where the incoming flow is assumed undisturbed by the presence of the car. The grid distribution on the symmetry plane for Grid A is shown in Figure (4). The grid density is high around the car where the flow gradients can be large. For the same reason, there are also densely packed meshes in the clearance region between the underside of the car and the ground. Figure (5a) provides a combined isometric view of the computational mesh on both the upper surface of the car and the symmetry plane. Figure (5b) shows a cross-sectional cut (in the z-direction) of the three-dimensional volume grid. There are 131,079 nodes and 727,890 tetrahedral cells in grid A. Grid B was generated by the simple stretching method in Gambit, which resulted in a coarser mesh around the car than Grid A. Grid B has 146,112 nodes and 673,483 tetrahedral cells. It is found that the size function used in generating Grid A can effectively wrap around the car body with much densely distributed grid points than the stretching scheme used for Grid B. A third Grid C was also generated by using the same size function as that used for Grid A. The distance of size function in Grid C is about one half of that used in generating Grid A. Hence Grid C has a much coarser grid distribution near the wind tunnel floor than that of Grid A. There are 134,095 nodes and 738,882 tetrahedral cells for Grid C. The major part of the results presented in this paper has been obtained by using grid A.

The resulting mesh file was then imported into FLUENT. In FLUENT, the three-dimensional, segregated, implicit steady solver was used. The velocity inlet and pressure outlet boundary conditions were used at the inlet and the exit surfaces that are shown in Figure (3), respectively. The no-slip wall conditions were applied on all car surfaces and on the floor of the wind tunnel denoted by box wall 3. Symmetric conditions were prescribed for both the symmetry plane and the surfaces of box wall 1 and box wall 2. No turbulence model was invoked in the calculations. In addition to the residuals of the numerically solved governing equations, we have also monitored the lift coefficient and the drag coefficient to determine the convergence of the computations. The simulation results
were presented in terms of velocity magnitude and gauge pressure. FLUENT uses the decomposition of the absolute pressure into the operating part and the gauge part to reduce the round-off errors when calculating low-speed flows; hence the pressure values presented in this report are all gauge values. The operating pressure has been set equal to the atmospheric pressure.

In the simulations, the test model was attached to the floor at the bottom of the wheel cover with a ground clearance of 1.6 inch. The air at room temperature has an incoming airspeed of 2.0 miles/hour. The Reynolds number based on the longitudinal length of the car is in the order of. This corresponds to the lowest Reynolds number used in the wind tunnel testing. The simulations have been performed on a laptop computer with a 2.2 GHz Pentium 4 CPU and 640 MB memory. The run time is normally around 7 hours for roughly 600 iterations.

SIMULATION RESULTS

The first-order upwinding method has been used in all the three grids and the numerical solutions were found to have converged for all of them. The less numerically dissipative second-order upwinding method in FLUENT was also activated for a simulation using Grid A. The evolutions of the residuals for the run with the second-order upwinding method are shown in Figure (6). The numerical simulation took about 600 iterations for the residual of the continuity equation to drop by about five orders of magnitude and was therefore considered converged.

The numerical lift coefficient evolutions during the iterations are shown in Figure (7) for all the cases calculated. The values of the calculated lift coefficients level out after about 400 iterations in all the calculations performed, although the calculations continue until all the residuals of the equations have dropped by five orders of magnitude. The measured lift coefficient is also included in Figure (7) for comparison. The lift coefficients have been normalized by the free stream dynamic pressure and the frontal area. The negative value of means that a downward force toward the ground is exerted on the car at this speed, which improves the cornering capability of the car [5]. The converged calculated lift coefficients using Grids A, B, and C compare well with the measured value. The differences between the numerical results and the measured data range from 4.7% for Grid A to 10% for Grid B. The result obtained by using the second-order upwinding method differs little from that using the first-order upwinding. The reasonably good agreement between the various calculated lift coefficients and the measured value indicates that the pressure field, which is the main contributor to lift for the current flow condition, has been adequately resolved.

Figure (8) shows a similar comparison of the drag coefficient. The cases and the notations are the same as those for Figure (7). The drag coefficient evolutions show that the calculated drag forces reach stationary values at nearly the same pace as those for the lift. The converged values, however, vary. The Grid A and the Grid C results agree very well with the measured value of 0.24. The result of the second-order scheme is somewhat lower than that of the first-order scheme. For Grid B, which has the coarsest surface meshes among the three grids used, the calculated drag coefficient deviates significantly from the measured value. The variation in the drag calculations points to a need to examine the calculated boundary layer on the car surfaces. A careful examination of the grids and the flow field solutions shows that, over some parts of the vehicle surfaces, the boundary layers are poorly resolved with only a few grid points inside the viscous dominated regions.

The pressure distributions on the surfaces of the car have been examined and presented in Figure (9). The regions of high pressure appear at the leading edge and the front wheel casings of the car. The similar high-pressure area can also be spotted near the canopy, which extrudes from the upper surface of the car. The airflow experiences resistance at these locations, hence it is expected that the pressure increases when velocity drops. From Figure (9), we can see that the flow field around the whole car surfaces develops in a relatively smooth manner except for the leading edge, the front wheel casing and the bulged structure under the car, which was designed to install the driver’s seat.
Figure (10) shows the velocity magnitude contours, normalized by the free stream value, on the cross-sectional plane shown in Figure (5b). The contour lines near the bottom wall of the wind tunnel show a boundary layer behavior in the near wall flow region. There is also a high-speed region above the canopy.

Another view of the velocity magnitude distribution is shown for the symmetric plane in Figure (11). The air flows from the left to the right parallel to the lower wall of the wind tunnel. The air moves along the surface of the car smoothly except nearing the region where the canopy connects to the car body, where low-speed regions are observed. The tail part of the car ends with a cusp and the flow from the upper and the lower surfaces merge smoothly. The velocity vectors and the streamline patterns have not shown flow separation in the region immediately downstream of the car body. The gage pressure (the difference between the absolute and the atmospheric pressure) contours on the symmetry plane are shown in Figure (12). Comparing with Figure (11), the highest pressure appears at the leading edge of the car where the velocity magnitude has its minimum. The lowest pressure area is on the top of the canopy where the air picks up the highest speed. The calculated quantities are consistent with the expected flow physics.

CONCLUDING REMARKS

For the initial simulations presented, we have successfully built the surface meshing of the Sunseeker test model. Converged numerical solutions for the flow field around the Sunseeker have been obtained. The calculated flow fields have been shown for three different grids and two numerical methods of different order of accuracy.

The computed lift coefficients agree well with the measured data for all the three grids. This agreement suggests that the outer, inviscid part of the flow around the Sunseeker has been adequately resolved by each and every one of the grids used and the results are accurate. The predicted drag coefficient values are less uniform, although the agreement is reasonably well for two of the cases calculated. To improve the drag force predictions, more computational cells are likely to be needed within the boundary layers that are expected to develop on the surface of the car. This measure can easily push the total number of cells over one million. Since the thickness of the boundary layer decreases with the increase of the Reynolds number, such refinements of the grid distributions in the boundary layer will become important for the higher car speeds. It is more so when transition or turbulence models are used at these high Reynolds numbers.

It should also be noted that the clearance between the underbody of the car and the wind-tunnel floor was not recorded during the testing that was carried out at the Applied Aerodynamics Laboratory of WMU. In the numerical calculations, the wheel cases were assumed to be 1.6 in above the tunnel wall. Since the two sides of the car extend toward the ground to accommodate the wheels in its original design, the spacing between the lower side of the main body and the side walls actually form an air passage that handles a significantly more amount of airflow than what one normally would see without the drop-down “side curtain”. To examine the effect of the clearance, we have run a case with the car sitting higher above the wind tunnel floor. The resulting calculated lift and drag coefficient were an order of magnitude smaller than the measured values. It can be argued that the ground clearance has some level of influence on the aerodynamic characteristics of the Sunseeker and needs to be modeled in an appropriate way.

Acknowledgments

The authors would like to thank J. Jacob for his valuable support in the Pro/E applications and A. Poot for his kind assistance in access to the Sunseeker project laboratory.

References


Figure 1. The 2003 Sunseeker and the team members. (Courtesy of Cathy Smith)

Figure 2. The experimental model of the Sunseeker sits in the test section of the WMU wind tunnel.
Figure 3. Configuration of the computational domain.

Figure 4. The tetrahedral mesh on the symmetry plane.
Figure 5a.

Figure 5b.

Figure 5. The tetrahedral computational mesh. (a) On the car surface and the symmetry plane. (b) On a cross-sectional plane.
Figure 6. The variation of the residuals of the numerical solved continuity equation and the momentum equations.

Figure 7. Convergence history of.
Figure 8. Convergence history of $C_d$.

Figure 9. Different views of the surface pressure distribution on the car surfaces.
Figure 10. Velocity magnitude contours on a cross-section plane.

Figure 11. Velocity magnitude contours on the symmetry plane.
Figure 12. Pressure contours on the symmetry plane.
Testing the Tests: An Investigation into the Effectiveness of Alternative Assessment Methods for Bilingual Language-Impaired Children

LENA G. CAESAR, Ed.D., CCC-SLP
Department of Educational Studies
Western Michigan University

Abstract

Alternative language assessment procedures are gaining in prominence as authentic alternatives to traditional measures, but scant data exist regarding the relative effectiveness of the data they provide. This paper first summarizes the criteria for effective language assessment of bilingual children as documented in the literature, and then comparatively evaluates the effectiveness of research studies employing the three most-cited alternative assessment approaches: descriptive, dynamic, and curriculum-based. Conclusions regarding the impact of appropriate assessment on issues of over- and under-representation of bilingual Hispanic students are also discussed.

Limited data exist regarding the impact of limited English proficiency on the referral and placement of bilingual children in special education and speech-language programs. The need for this type of information is becoming increasingly urgent. Minority populations are growing at unprecedented rates and the nation’s classrooms are becoming rapidly multicultural. Of the 63 racial/cultural categories identified by the US Bureau of the Census (2000), Hispanics comprise the fastest growing minority population in the United States —having grown from about 9% of the United States population in 1990, to about 13% in 2000. Whereas in 1982, Hispanics comprised only 1 in 10 school-aged children, this number is expected to increase to 1 in 3 by the year 2020 (Gersten & Woodward, 1994; US Census Report, 2000).

The assessments of children who are linguistically different as well as culturally diverse has become one of education’s greatest challenges. The stakes multiply in importance for students being assessed for special education or speech-language pathology (SLP) services. For such students, the assessment process has been called into question, especially in light of the fact that students from minority populations (especially Hispanics and African-Americans) are disproportionately overrepresented in special education. For example, according to data cited by Correa and Heward (2000), although 32.5% of the entire school-age population can be classified as minority, nearly 40% of children identified as mild or moderately mentally impaired belong to a minority group. In terms of the Hispanic population, evidence of both over- and under-representation exists. Current data indicate that 80% of Hispanics referred for special education services are placed in learning disability (LD) and speech-language pathology (SLP) programs— both of which directly involve language proficiency and competence. The rise in the number of Hispanic students now being referred to LD and SLP programs has been accompanied by a comparable reduction in the number of Hispanic students now being placed in programs for children who are mildly mentally retarded or gifted and talented (Ortiz, 1997).

Several possible causes for this disparity have been cited in the literature. Burnett (2000) attributes the crisis to the national scarcity of qualified bilingual personnel and the lack of appropriate assessment tools, but Ortiz and Garcia (1995) place the blame on a broader set of causative factors including

the absence of guidelines and data which can be used to develop procedures which (a) help distinguish cultural or linguistic differences from disabilities, (b) yield a non-biased assessment, (c) assure due process in decision making, and (d) result in individualized education programs which help language-minority students with disabilities achieve their potential (p.147).
Most teachers and special education personnel have not been trained to determine differentially whether the academic problems that many bilingual students experience stem from a language learning disability or a language difference associated with typical second language acquisition (Gersten & Woodward, 1994). As a result, many students are misidentified and misplaced in special education programs with educators who are ill-equipped to address academic issues related to second language learning (Gersten & Woodward, 1994; Jitendra & Rohena-Diaz, 1996).

The issue of non-biased assessment for students who are culturally and linguistically diverse has received adequate attention in both legislation and litigation. As early as 1964, Title IV of the Civil Rights Act required that language assessment of children with limited English proficiency be conducted in both the native language and English. This principle was subsequently re-affirmed by both Section 504 of the Rehabilitation Act of 1973 and the Education for All Handicapped Children Act of 1974, subsequently renamed the Individuals with Disabilities Education Act (IDEA) in 1990. The latest (1997) IDEA amendments revisited the issue by specifying that assessment for the purpose of identifying and placing children with disabilities should be conducted in the child’s native language.

In spite of adequate legal basis, the reality of non-biased assessment has not been easy to attain in any of the three areas typically assessed by special services personnel—psychological, language, or educational. Traditionally, standardized, norm-referenced measures have been used to obtain much of this information, even though such measures involve testing primarily in English, are norm-referenced on monolingual English speakers, and are incapable of assessing the student’s relative English and native language proficiency (Jitendra, 1995). The problem is compounded by the fact that few tests exist in Spanish, and of the few available, most are translated versions of English tests. Such translations contain items that are (a) culturally inappropriate, (b) linguistically incorrect, or (c) contextually irrelevant (Maldonado, 2000). Jitendra (1995), in her discussion of limitations of traditional ‘discrete point’ language assessment measures, categorizes such tests as “spurious, decontextualized, inadequate, instructionally aloof and thus unsuitable for making decisions regarding eligibility, placement, and instructional decision making” (p. 45). In the specific context of bilingual proficiency, standardized tests not only are inadequate in terms of reliability and norming but also biased in their ability to provide information capable of distinguishing a language learning disability from a language difference (Ortiz, 1997).

If the situation is to improve, qualitative changes are needed in the assessment of bilingual students. Burnette (2000) specifies in her list of solutions and best practices the principle of individualized assessment that takes into account not only the child but also his/her environment. Burnett recommends that traditional instruments be combined with qualitative assessment measures from a variety of sources (such as observations and interviews) and from a variety of settings (school, home, community).

Questions remain about the range of alternative assessment procedures that might be used to conduct accurate and non-biased assessment of bilingual students with suspected language learning disabilities. This paper seeks to shed light on this issue by (a) summarizing the characteristics of effective alternative bilingual language assessment, as documented in the literature; (b) evaluating the effectiveness of research studies employing alternative language assessment measures, and (c) drawing conclusions regarding the impact of appropriate assessment on issues of over- and under-representation of bilingual Hispanic students in special education and related services programs.

Review of Literature

Criteria for Effective Bilingual Assessment

For the purposes of analyzing the effectiveness of the assessment measures employed in the literature being reviewed, five criteria were selected as essential for determining assessment effectiveness. These were drawn from the general literature on alternative language testing related to bilingual language assessment of children with special needs (Baca, 1994; Barrera, 1995; Burnette, 2000; Jitendra, Rohena-Diaz, & Nolet, 1998; Ortiz & Garcia, 1995). Alternative assessment measures were evaluated as to whether they conformed to each of the criteria as outlined below:

1. Native Language and Classroom Language Assessment (L1 & L2). Measures employed /discussed in the document should be capable of assessing both the child’s native language (L1) and the classroom language (L2). Ortiz & Garcia (1995) stated clearly that “every language minority child referred to special education should receive a comprehensive language assessment in his or her native language and in English” (p. 475). Baca (1990) described
the issue of ‘English-only testing’ of language minority students as perhaps the fundamental cause of minority over-representation in special education programs.

2. Comprehensive Language Assessment (CLA). Procedures described/discussed in the document should provide in-depth information regarding language performance. According to Burnette (2000), a description of “in-depth performance” is based on whether the assessment measure allows the evaluator to obtain descriptive information from a variety of sources and in a variety of environments.

3. Language Dominance Assessment (LDA). The approach outlined in the document should specifically address the issue of language dominance. Language dominance, as defined by Jitendra (1996), “is the language most often used by the student for communication and self-expression” (p.43). Ortiz (1997) expanded the concept when she described dominant language as

the language the student: (a) first learned, (b) prefers to use, (c) consistently chooses to use when speaking with bilingual individuals who speak the same dialect, and/or (d) shows the greatest ease in using. It is also the language that seems to have a greater influence on the other language. (p. 43)

4. Language Proficiency Assessment (LPA). Procedures described/discussed in the document should assess relative language proficiency. Language proficiency can be defined as the level of competency an individual demonstrates in a particular language. Ortiz (1997) listed four distinguishing characteristics of a language proficient individual as (a) the ability to understand messages even if distorted; (b) the ability to express meaning clearly; (c) the ability to adequately use language in a variety of settings; and (d) the ability to self-correct.

5. Intervention and Placement Information (IPI). The results (outcomes) of the specific alternative assessment procedures should provide information regarding language learning capacity and appropriate placement. A mere description of students’ language performance is insufficient to provide adequate information for placement purposes. Results should reflect the fact that effective assessment should always lead to appropriate placement (Baca, 1990).

Organizational Schema

The research literature contains a variety of terms referring to non-standardized assessment approaches. These include “naturalistic assessment” (Baca, 1990), “alternative assessment” “criterion-referenced assessment”(Ruiz, 1995), “informal assessment” (Burnett, 2000) and “non-psychometric or non-standardized assessment” (Baca,1990). For the purposes of this study, the term “alternative” is favored as a means of avoiding the ambiguity produced by the subtle definitional variations of the other terms. The “non-standardized” nature of these approaches has given rise to a quantity and variety of procedures that are virtually unlimited. However, the three most frequently cited alternative approaches in the bilingual special education literature—from descriptive, (b) dynamic, and (c) curriculum-based approaches—comprise the categories addressed in this review (Baca, 1990; Cline, 1998; Jitendra, 1996; Olswang & Bain, 1996). For purposes of specificity, the broad category of descriptive approaches is further subdivided to accommodate two types of descriptive language assessment procedures: (a) language sample analysis and (b) observational/rating scales. What follows is a description/definition of each of these approaches as outlined in the literature.

Descriptive Approaches. The goal of the descriptive assessment approach to bilingual assessment, as articulated by Damico (1991), is to “collect data that are meaning-based and integrative” (p.179). Thus this approach (frequently referred to as a ‘naturalistic’ approach) aims to describe and analyze communication as it occurs across languages, in a variety of naturalistic contexts and interactive partners. In the context of bilingual assessment, this approach is not only capable of assessing both native and second language competence but is also useful for providing information on the individual’s relative language proficiency (Jitendra, 2000).

Procedures employed by the naturalistic approach are numerous, highly individualized, and varied. In the area of language assessment, three main types of descriptive approaches predominate: (a) language sampling procedures, (b) rating scales and protocols, and (c) direct observation (Jitendra, 2000). In this paper, studies employing descriptive approaches are re-grouped and sub-classified according to two areas only: (a) elicited and spontaneous language sampling and (b) observational and rating scales.

Dynamic Assessment. The concept of dynamic assessment was initially described by Vygotsky as part of his
model of cognitive development. Vygotsky (1986) proposed that a child’s knowledge develops within a “zone of proximal development” (ZPD), as experiences are mediated and shared with more capable partners. In the context of language assessment, dynamic assessment can be defined as “an instructionally-oriented model of assessment that serves the dual purpose of accurately identifying a student’s instructional language needs and planning instruction” (Jitendra, Rohena-Diaz, & Nolet, 1998, p.182).

Assessment procedures most frequently used for the dynamic assessment of language include (a) testing-the-limits, (b) graduated-prompting and (c) test-teach-retest. All dynamic assessment approaches “incorporate a learning component into the testing situation and examine the learner’s responsiveness to teaching” (Jitendra, Rohena-Diaz, & Nolet, 1998, p.182). Whereas testing-the-limits and graduated-prompting approaches are useful for determining readiness for and progress in intervention, test-teach-retest procedures are better suited for distinguishing language differences from language learning disabilities (Guittierrez-Clellen & Pena, 2001). Test-teach-retest procedures are designed to test children learning potential by providing them with mediated learning experiences (MLE) following initial testing and then retesting at the end of instruction. The idea is that a child with a typical language-learning system (or without a disability) will be able to benefit immediately from instruction, whereas the student with a disability will have difficulty learning even when explicit instruction is provided.

Curriculum-Based Language Assessment (CBLA). Curriculum-based language assessment measures are designed to first identify the areas within the curriculum where language-related problems are most likely to occur and then to assess language performance using the context and content of the curriculum. Contrary to popular notions, the term “curriculum” is not limited to the design of specific academic courses. Nelson (1994) broadly defined curriculum as “the variety of things children are expected to learn in school to become successful, independent citizens” (p.105). Nelson’s description of the six kinds of curricula most often observed in schools includes both explicit academic-related types of curricula and the more implicit, subtle types related to social expectation and communicative interaction. These are (a) the official curriculum defined by school districts; (b) the cultural curriculum dictated by the unspoken expectations of the mainstream culture; (c) the de facto curriculum governed by textbook selection in individual schools/classrooms; (d) the school culture curriculum determined by both the stated and the unstated rules that determine acceptable classroom behavior; (e) the hidden curriculum controlled by teachers’ conscious or unconscious values for desirable or undesirable classroom behavior; and (f) the underground curriculum dictated by peer-determined rules about acceptable and unacceptable age-appropriate social behaviors.

Curriculum-based language assessment (CBLA) is distinctly different from more generalized curriculum-based assessment or curriculum-based measurement that seeks to determine the student’s instructional needs through an analysis of the local curriculum. In contrast, CBLA looks at the types of language skills and strategies the student uses during all school-related communication breakdowns, assesses their effectiveness, and then determines the types of curricular and student adaptations that may be necessary to communicative success (Nelson, 1994; Schoenbrodt, Kumin, & Sloan, 1997). Baca & Valenzuela (1994) go so far as to posit that CBLA, which employs “criterion-referenced, informal and teacher-made devices” should be “the first step” (p. 5) in the assessment process of bilingual children with disabilities. This information, they believe, is crucial to the appropriate instructional placement of the student.

Results

Language Sampling

Nine of the documents reviewed utilized or recommended language sampling procedures as valid alternatives to traditional testing for bilingual children. Seven of these studies were experimental documents describing research conducted with Spanish-speaking subjects, whereas two of the studies were opinion-based. Studies also varied in terms of method of elicitation (spontaneous versus elicited) and sample type (conversational versus narrative).

Two of the nine experimental studies sampled the narrative or story-telling abilities of their subjects. Goldstein et al. (1993) employed a standardized story retell task with four-scene cue cards to elicit the students’ narratives, and Guittierrez-Clellen (1994) analyzed syntactic complexity in spontaneous narratives.

The other experimenter with the exception of Bedore & Leonard (2001) who used structured elicitation tasks, utilized spontaneous, conversational samples in their assessment protocols. Many of the studies, however, focused on analyzing only one aspect of subjects’ linguistic competence, generally syntactic/grammatical structures.
These structures included article use, article + noun agreement, morphosyntactic structures, and grammatical morphemes. Although “language samples and story telling procedures can be used effectively to establish a relative language proficiency profile” (Barrera, 1995, p.8), the majority of the language samples were obtained in English or Spanish, not both and thus were not capable of providing information on subjects’ relative language proficiency or language dominance. Only one study (Langdon, 1999) reported assessing subjects in naturalistic contexts with a variety of conversationalists in both English and Spanish. Therefore, this study (Langdon, 1999) was the only document that conformed to all six measures of effective alternative assessment that this paper specified. Table 1 shows the degree to which the language sampling methods met the 6 effectiveness criteria.

Observational/Rating Scales. In the category of observational/rating scales, seven studies were reviewed for evidence of effective assessment using the five pre-determined criteria. All of the studies utilized parental reports or interviews to obtain their data. Three of them (Gonzalez, 1994; Jackson-Maldonado, 1999; Restrepo, 1998) coupled parent reports with teacher observation and rating scales. Five of the studies (74%) reported obtaining information on students’ English and Spanish competence, language dominance and language proficiency.

Six (85%) reported that information obtained spanned a variety of communication contexts and a number of linguistic modes. All of the researchers found that parental reports provided valid data for differentiating between a language disability and linguistic differences attributable to second language learning. The one obvious weakness with this method was in the area of intervention or curricular relevance. None of the studies addressed the issue of how the information obtained via this procedure could generalize to the classroom/intervention setting. Table 2 summarizes the results of this evaluation.

Dynamic Assessment. Three of the six (50%) documents reviewed (Butler, 1997; Jitendra & Rohena-Diaz, 1996; Jitendra, Rohena-Diaz & Nolet, 1998) were opinion-based, literature review-type articles and focused on the strengths and advantages of dynamic assessment (DA) over more traditional standardized testing. Both of Jitendra’s articles employed hypothetical case studies as a means of illustrating the ideal dynamic assessment process. In both articles, subjects for the case study were 8 to 10-year old males of Puerto Rican descent for whom Spanish was considered the home language. These studies demonstrated the ability of DA to differentiate effectively between language disability and linguistic differences. Although Butler’s tutorial did not employ the case study approach, she also discussed the benefits of employing DA methods for testing children who are culturally and linguistically diverse.

The other three articles were descriptions of experimental studies in which the validity of DA as an alternative assessment measure was explored. All of these studies were done by Pena and colleagues and focused on using DA as a supplemental assessment measure for bilingual children of Hispanic descent. The majority of the subjects in all the studies were preschool-aged children, and all were enrolled in bilingual Head Start programs. The three experimental studies employed standardized tests for the pre-test portion of the assessment protocol, and all six documents utilized or supported a test-teach-test design that included a mediated learning experience (MLE), during which children were exposed to tasks that were similar (but not identical) to tasks presented during the pretest. Because most of the subjects were bilingual, responses were accepted in both English and Spanish.

An analysis of the six documents reviewed, indicated that DA complied with all of the six criteria for effective assessment as stipulated in this paper. All the studies reviewed (a) focused on Spanish-speaking children at risk for learning disorders, (b) assessed more than one area of language, (c) were designed to address issues of language dominance and proficiency, (d) provided relevant intervention information and proved capable of distinguishing a language-learning disability from a language difference. Table 3 provides a summary of the effectiveness ratings of this method.

Curriculum-Based Language Assessment. The application of CBLA approaches to the language assessment of bilingual students has not received extensive coverage in the literature. Only five candidate documents met criteria for inclusion in this review. In four out of the five documents (80%), students were assessed in specific language-related curricular areas. These included (a) reading comprehension and fluency (Baker & Good, 1994), (b) language arts activities based on story passages (Duran & Szymanski, 1994), (c) Head Start curricular-based vocabulary (Steffani, 1992), and (d) storybook writing (Ruiz, 1995). The subjects of all four studies were Spanish-English bilingual students, and assessments were carried out in both languages by bilingual personnel. The fifth document (Cline, 1998), was opinion based and focused on the importance of not only assessing the child in his/her
The fourth document (Ruiz, 1995) was somewhat singular in its approach and utilized a classroom-based observation as a means of determining the validity of the theory that the performance of children in special education is affected by their interactional/instructional contexts. The Ruiz (1995) study, along with the documents discussed previously, all (100%) utilized materials that were directly linked to the students’ academic context and thus met the criteria for instructional relevance. However, only two documents (Cline, 1998; Ruiz, 1995) dealt with or discussed children with disabilities. Table 4 summarizes the results of the analysis.

Discussion

In this review of the literature, three types of alternative assessment measures—descriptive, dynamic, and curriculum-based—were analyzed in terms of their effectiveness in providing non-biased assessment information for at-risk bilingual Hispanic students. Each document was first classified according to either a descriptive, dynamic or curriculum-based approach and then evaluated on the basis of five specific criteria gleaned from the literature on bilingual assessment. Although documents were individually assessed, the focus of this paper was not so much on whether specific procedures met criteria but whether the approaches were capable of effectively assessing bilingual Hispanic students.

Interestingly, findings indicate that regardless of approach, only 20% of the individual documents met all the assessment effectiveness criteria specified in this paper. Eighty percent of these “one hundred percent criteria” articles were classified in either the dynamic approach group or the curriculum-based language assessment group. However, dynamic assessment approaches were found to more fully comply with the criteria for effective assessment than curriculum-based language assessment approaches. Descriptive measures, whether employing language sampling procedures or observation/rating scales, were found to be the least capable of conforming to the efficacy criteria as outlined in this paper.

This analysis also served to highlight the strengths and weaknesses of each approach. For example, whereas 100% of the documents in the DA and CBLA classifications met the instructional and placement information (IPI) criterion, none of the documents in the descriptive classification of language sampling or observational/rating scales met this criterion. However, descriptive measures do have the capacity to obtain spontaneous information regarding language proficiency and dominance in ways that DA approaches cannot. Documents in the category of observational rating scales also more consistently provided comprehensive language data (i.e. information from a variety of contexts and language components) than did both DA and CBLA approaches. The findings on each approach’s relative strengths and weaknesses may suggest that an approach’s failure to conform to all of the predetermined criteria for effective assessment may not necessarily disqualify it as a useful assessment measure.

The results provide strong evidence of variation among alternative assessment approaches. That is, not all are equally effective in their ability to assess bilingual Hispanic students. Furthermore, their effectiveness may be affected by the choices in administration procedures selected by the examiner. For example, although language sampling is clearly an alternative measure capable of being administered in the student’s first and second languages, less than 50% of the documents reviewed, assessed or recommended assessment in both languages. Further, data from this study also suggest that measures that purport to provide instructional and placement information may not comply with all the basic criteria that the literature recommends. This finding may suggest that a combination of alternative and traditional approaches, rather than any single approach, may be most effective in assessing the language of bilingual students.

This study also exposed several inadequacies in this area of research. First, it highlighted the need for more data-based research on alternative assessment of bilingual Hispanic children. This was demonstrated by the meager 26 documents that met criteria for this evaluation, despite the fact that the search spanned a 10-year period, and only 80% of these were data-based. Secondly, even those that were data-based involved small numbers of subjects who were of a limited age span (usually preschool). Thirdly, most of the existing research literature was authored by a limited number of researchers whose names recurred multiple times in the studies analyzed. The apparent lack of research documenting alternative assessment efficacy may suggest the absence of widespread usage. Given the increasing numbers of bilingual children being placed in special education programs and the documented inability of standardized language measures to adequately assess bilingual children, the fair and appropriate placement of
linguistically diverse students may be legitimately challenged (Gersten & Woodward, 1994).

Other limitations inherent in this type of research include (a) the difficulty of ascertaining that all the relevant documents have been identified and reviewed, (b) the fact that this review was limited to documents dealing with language assessment only, to the exclusion of psychological and educational assessment, and (c) the reality that this research was conducted in an area where the lack of scientific evidence requires the researcher to draw subjective and inferential conclusions regarding assessment effectiveness.

These limitations, however, may be useful in determining new pathways for future research. More experimental research is definitely needed with culturally and linguistically diverse students. Specifically, there is need for data on specific strategies for effectively assessing children from Hispanic backgrounds—the nation’s largest minority group. Presently, there is little research providing follow-up information on the academic success of students who have been tested by alternative methods. The crisis of minority over-representation in special education, may actually stem not so much from a dearth of assessment approaches but from an absence of rigorous scientific grounding that validates assessment decisions. Perhaps it is time to test the tests—and thus guarantee appropriate education to all the nation’s children.

References


### Table 1

**Effectiveness Rating for Language Sampling**

<table>
<thead>
<tr>
<th>Study</th>
<th>L1 &amp; L2</th>
<th>CLA</th>
<th>LDA</th>
<th>LPA</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedore &amp; Leonard (2001)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Eng &amp; O’Connor (2000)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Goldstein et al. (2001)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Gutierez-Clellen (2000)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Jackson-Maldonado (1990)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Langdon (1999)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Restrepo &amp; Kruth (2000)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Restrepo &amp; Gutierez-Clellen (2001)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Restrepo &amp; Silverman (2001)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

**Note.** L1 = 1st Language; L2 = 2nd Language; CLA = Comprehensive Language Assessment; LDA = Language Dominance Assessment; LPA = Language Proficiency Assessment; IPI = Intervention & Placement Information.

### Table 2

**Effectiveness Rating for Observational/Rating Scales**

<table>
<thead>
<tr>
<th>Study</th>
<th>L1 &amp; L2</th>
<th>CLA</th>
<th>LDA</th>
<th>LPA</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonzalex (1994)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Gonzalez et al. (1996)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Jackson-Maldonado (1990)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Patterson (2000)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Restrepo (1998)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Thal et al. (2000)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Note.** L1 = 1st Language; L2 = 2nd Language; CLA = Comprehensive Language Assessment; LDA = Language Dominance Assessment; LPA = Language Proficiency Assessment; IPI = Intervention & Placement Information.

### Table 3

**Effectiveness Rating for Dynamic Assessment**

<table>
<thead>
<tr>
<th>Study</th>
<th>L1 &amp; L2</th>
<th>CLA</th>
<th>LDA</th>
<th>LPA</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutierrez-Clellen (2001)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Jitendra &amp; (Rohena-Diaz 1996)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Jitendra et al. (1998)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lidz &amp; Pena (1996)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pena &amp; Quinn (1992)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pena et al. (2001)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**Note.** L1 = 1st Language; L2 = 2nd Language; CLA = Comprehensive Language Assessment; LDA = Language Dominance Assessment
Assessment; LPA= Language Proficiency Assessment; IPI=Intervention & Placement Information.

Table 4
Effectiveness Rating for Curriculum-Based Language Assessment

<table>
<thead>
<tr>
<th>Studies</th>
<th>L1 &amp; L2</th>
<th>CLA</th>
<th>LDA</th>
<th>LPA</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baker &amp; Good (1994)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Cline (1998)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Duran &amp; Szymanski (2000)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ruiz (1995)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Steffani (2000)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Note.  L1 = 1st Language; L2 = 2nd Language; CLA = Comprehensive Language Assessment; LDA= Language Dominance Assessment; LPA= Language Proficiency Assessment; IPI= Intervention & Placement Information