Informatics in Education, 2006, Vol. 5, No. 1, 1–16

Gender and Programming Contests: Mitigating Exclusionary Practices

Maryanne FISHER
Department of Psychology, Saint Mary's University
923 Robie Street, Halifax, Nova Scotia, B3H 3C3, Canada
e-mail: mlfisher@smu.ca

Anthony COX
Faculty of Computer Science, Dalhousie University
6050 University Avenue, Halifax, Nova Scotia, B3H 1W5, Canada
e-mail: amcox@cs.dal.ca

Received: January 2006

Abstract. Individuals vary across many dimensions due to the effects of gender-based, personality, and cultural differences. Consequently, programming contests with a limited and restrictive structure (e.g., scoring system, questioning style) are most favourable and attractive to a specific set of individuals with the characteristics that best match this structure. We suggest that a more inclusive and flexible structure will allow contests to be more appealing to a wider range of participants by being less biased towards specific traits. As well, by making contests more broadly appealing, they become better post secondary recruiting tools that can potentially be used to attract under-represented populations to the discipline of computer science. In this paper, we focus on gender-based differences and the effect of a competition's structure on female participants.

Key words: programming contests, scoring systems, gender bias.

1. Introduction

It is well known that females are equally capable to males, but less represented in computer-oriented professions (Cooper and Weaver, 2003). This under-representation is the result of many interacting gender-based, social, and cultural factors. While it is beyond the scope of this paper to examine these factors in detail, we focus on one element, gender, as it leads to several promising and novel ideas. Specifically, we suggest that programming contests can become more inclusive to female participants if changes are introduced that make the contest more sensitive to gender-based differences. This paper begins by examining these differences from a theoretical perspective, using relevant literature in psychology, sociology and education. Then, we examine specific contest components, such as scoring, which could be improved to be more inclusive and less biased. It should be stated that, although this paper focuses on gender, the results are by no means limited to this variable alone, as Schofield (1995) found that many of the factors that dis-

infon v.1998/05/11 Prn:17/03/2006; 14:07 F:INFE073.tex; VTEX/RRR p. 2

suade female use of a U.S. high-school computer center are equally applicable to cultural groups such as African Americans.

Fully inclusive contests that are enticing to a wide range of participants provide computer science departments with a high profile tool for advertising the discipline and potentially addressing issues of under-representation by specific population groups. Actively attracting students to the discipline by providing an enjoyable, seemingly fair and unbiased recreational event can lead to increased interest, and hence provides a recruiting tool for post-secondary education. With respect to gender-based under representation, it is known that women will engage in competitive activities (Hrdy, 1999), but will be most likely to do so when they believe they can succeed in these activities. By providing a contest environment that does not inherently disadvantage women, as we suggest some current practices do, it is possible that female participants will be more successful at International Olympiad in Informatics (IOI)[1] level competitions, and thus, be more likely to pursue careers and educational opportunities in computing and information technology.

We are not suggesting that this paper presents a solution to the problem of female under-representation in computer science or their low level of participation in programming contests. Rather, the issues we examine and the practices we suggest in response to these issues should be viewed as an element of a larger solution strategy. While changing a contest environment can minimize the likelihood of gender-based disparity occurring during competition, it will not address the complex social and cultural factors that dissuade women from participating. However, by avoiding exclusionary practices and providing a fair and equitable environment, computing contests can be part of the solution as opposed to perpetuating the problem.

This paper will now review some of the literature pertaining to gender-based differences in cognition and behaviour that affect performance in competitive events. We will review different elements of competition that are particularly gender sensitive, under a conceptual framework that considers current research findings to be applicable to programming contests. This examination is followed by some suggestions that, based on the reported differences, will lead to a fairer, more sensitive, more inclusive and less biased competitive environment.

2. Factors Influencing Individual Performance

We believe that gender influences how people will perform within computing contests, although it is presumable that there are numerous other variables, such as ethnicity, personality, experience, and so on. We focus on gender because there exists a large body of contemporary research on the interaction of gender and competition, whereas current literature is more limited with respect to these other variables. In addition, casual observations, as well as various reports (e.g., Helgeson, 2005), reveals a systemic under-representation of women in computer science. We do not address the reasons for low institutional enrollment, and instead, our contribution will be to suggest reasons why

[1] See, `http://www.ioinformatics.org`, for additional detail.

women may feel excluded from these contests, or are prevented from maximizing their performance. To address these issues, we identify practices that will lead to higher levels of participation and success. Correspondingly, increased participation can be used to improve enrollment when computing contests are used as a vehicle for advertising the discipline and recruiting college-level students.

Given that these contests are competitive, we begin by discussing gender differences in perceptions and acceptance towards competitive environments. Previously, researchers have argued that women's low representation rates in competitions, such as in athletic contests, was simply due to women's cooperative nature. That is, it has long been argued that women are far more cooperative than men, and will shy away from any competitive situation in favour of reaching an amicable "everyone wins" solution. This belief is still widely held today, but is only correct within limited extents. Females are able to compete, and do so willingly, even when there is only one winner and a clearly ranked hierarchy of finishers. For example, a cursory view of the Scripps National Spelling Bee participation data, which is a voluntary contest held annually in the United States, readily shows that females are active in these contests at levels nearly equivalent to males (127 females and146 males participated in 2005). Women can compete as vigorously as men, and sometimes enjoy it as much as men (Hoyenga and Hoyenga, 1993).

Although females are able and willing to compete, and do so well in specific situations, the widely held notion that they tend to be more cooperative and less competitive than males does hold merit. In general, findings from social and developmental psychology indicate that females prefer cooperative activities, while boys prefer competitive activities (see (Hoyenga and Hoyenga, 1993) for a more comprehensive review). In competitive games, where the rules state that everyone must compete to win, males will often compete individually, while females will form small, cooperative alliances, with the group competing as a single entity (Hughes, 1988). Also, although public school teachers tend to use cooperative learning strategies less frequently than competitive strategies to aid student learning, cooperative strategies allow girls to achieve on par or higher than boys in math-related tasks (Fennema and Peterson, 1987). In fact, competitive games have been repeatedly found to give boys an edge, while cooperative activities generally give girls the advantage (Fennema and Peterson, 1987).

The problem is that, although women's performance is not *always* jeopardized by a competitive environment, exposure to these settings may be detrimental to women's performance at higher rates than for men (Hoyenga and Hoyenga, 1993). This contention should not be mistaken as an attempt to say that women are unable to compete, or that activities that are competitive in nature explicitly rule out women's participation; in contrast, we argue that these sorts of environments may simply disallow women to perform at their best. It must also be remembered that there is a stigma preventing women being positively viewed as competitive (Johnston and Crawford, 1999), and therefore, it is hardly surprising that women will sometimes label their behaviours as more cooperative then men's, when in fact it is equally competitive (Stockard *et al*., 1988). Moreover, women often feel pressured by popular stereotyping to be considered as passive and feminine, and thus actively hide their competitive nature (Cashdan, 1999). We now outline some of the variables that occur in competitive environments, which may lead to decreased performance by women.

2.1. *Risk*

When it comes to problem solving, four distinct approaches have been identified: apply a solution that is familiar, solve it via logical-mathematical reasoning, use trial and error whereby one works backwards from a possible answer and plugs in values or the use of a guess and check approach, and lastly, using a one-shot guess (Byrnes and Takahira, 1993). While all these strategies, and indeed the process of choosing a strategy, have an element of risk, the latter two strategies can be viewed as having higher levels of risk than the first two. Hence, the performance of risk averse individuals is more likely to be impacted when they cannot use the first two strategies. Thus, an interaction occurs between one's risk taking behaviour and contest questions that cannot be answered using familiar solutions or logical reasoning.

It has been well established among social psychologists that men are more risk prone, or take more risks, than women (see (Daly and Wilson, 2001) for a comprehensive review). Instead, women tend towards a strategy of risk aversion, and accept risk only as a last resort, or when the benefits are maximized and costs minimised. In general, women tend to choose high probability, low payoff strategies, whereas men tend to choose, under the same situations, low probably but high payoff strategies (Daly and Wilson, 2001). This difference in strategy may lead to differences in programming contest performance. We predict that men will be more likely to rush at forming a solution, completing it as quickly as possible, and then submit it for judging. Women will be more likely to devote more time to conceptualizing a solution, spend longer working on an individual solution, and then submit it for judging. This difference means that women will have less opportunity to try various solutions, in the belief that they have increased their accuracy by spending longer on a solution.

As a society, successful risk taking is often admired, and it seems that men are responsive to this possibility. Men frequently take greater risks in the presence of others, whereas this elevation appears absent among women (Daly and Wilson, 2001). Thus, within the current context, when contests involve teams, especially mixed-sex teams, we might expect that men will take increased risks with solutions, delegating team efforts to support these risks, while women will not feel comfortable with these investments. These effects, with respect to computing contests, will be elaborated upon later in the paper.

Furthermore, men tend to be increasingly risk prone when the need to succeed, or the fear of failure, is intensified (Daly and Wilson, 2001). This process is even greater when only one, high-risk behaviour seems plausible as other options represent seemingly dead ends (Daly and Wilson, 2001). Again, within the context of computing contests, where there are strict time limits and only one viable solution strategy, men will more rapidly identify a plausible solution in the hopes that it is "close enough" to being correct. Women will feel pressured in these situations, wanting to avoid unnecessary risk that consequently leads to feelings of anxiety, which for computing tasks is deleterious to performance (Cooper and Weaver, 2003).

In summary, we believe that the well researched gender difference in risk perception and acceptance is significant for a variety of reasons. First, accepting risks may be perceived as an indicator of confidence, and perhaps, competence (Daly and Wilson, 2001). Thus, within the context of programming contests, women may be perceived as less competent because they adopt less risky approaches. This matter is even more deleterious, given that stereotypical beliefs, such as "people view me as less competent because I am female" often lead directly to depressed performance (Steele and Aronson, 1995). Second, rather than gambling on their ability to guess the correct approach, women will take more time to explore one solution as opposed to implementing a "fast and dirty" solution that may not work. This less risky strategy will allow less time for trying a wider assortment of potential solutions, and less time to attempt remaining questions. Third, women will not guess at a response as readily as men, nor direct team mates to consider a possible solution, when they are not sure of themselves. They will also feel uncomfortable when, in a mixed-sex team, they are asked to pursue high-risk strategies.

2.2. *Pressures to Perform*

There are numerous pressures perceived by individuals within the context of competitions or contests. These pressures directly impact one's ability to perform through subsidiary mechanisms, such as the anxiety resulting from the pressure to perform. In this section, we examine some of the factors that influence an individual's perception of the pressures to perform.

2.2.1. *Time Constraints*

One area of cognition that we have been currently investigating is the spatial rotation of objects. While this task significantly differs from programming, it is a robust and frequently cited example of gender-based differences in performance and is therefore a useful predictor of similar differences in other situations.

Numerous studies (e.g., Voyer and Saunders, 2004) have documented a male advantage on paper-pencil tests where participants are asked to rotate a configuration of three-dimensional blocks and match the correct solutions to the rotated sample (Vandenburg and Kuse, 1978). We raise this issue because our research, as well as other published studies, indicates that the sex difference in this ability is an artifact of the measure. Thus, it is not the ability to perform mental rotation that differs between the sexes, but the impact of the measurement device (test) on males' and females' behaviour. For computing contests, where it is known that women have equal abilities, it is likely that a similar effect occurs and it is the contest environment, not female participants' abilities, that affects performance. Specifically, we have focused on two measurement factors, time constraints and negative scoring, as possible influences on women's ability to succeed in these environments. It is important to note that the issue of time constraint bias has also been documented with respect to standardized testing (Gallagher *et al*., 2000) and is not exclusive to mental rotation ability.

In the typical situation involving the mental rotation test, participants are given ten minutes to complete 24 questions. Usually, with this protocol, men perform at a higher

infon v.1998/05/11 Prn:17/03/2006; 14:07 F:INFE073.tex; VTEX/RRR p. 6

level than do women (Voyer and Saunders, 2004). However, using tight time restrictions promotes the use of risk assessment decision-making. For example, should one spend more time ensuring a few answers are correct, or devote less time to each question ensuring more are answered at the possible cost of accuracy? Women tend to adopt the first strategy, signaling the use of risk aversion, whereas men tend to adopt the latter strategy, suggesting risk proneness. Evidence for these choices is provided by Gallagher *et al*. (2000) who report that on the Scholastic Assessment Test Mathematics Section (SAT-M), a timed test, women leave more items blank (unattempted) than do men.

When there is no time limit and participants are simply asked to complete the mental rotation test, the sex difference in performance disappears, seemingly because women spend longer than do men in answering the test questions (Goldstein *et al*., 1990). In fact, recent findings (Voyer and Saunders, 2004) clearly show that men are far more prone to guessing responses on these tests than women, and women show a far greater reluctance to guess. Additionally, men tend to assume that their guesses will be correct, whereas women tend to assume that their guesses will be incorrect (Voyer and Saunders, 2004). During programming contests, we expect the same thought processes to be apparent, perhaps leading to lower performance by women.

2.2.2. *Scoring Concerns*

The issue of sex-differences in performance is further exacerbated by the use of negative scoring on the mental rotation test. The typical instructions for the test state that every incorrect answer will result in a score of negative one, while every correct response will result in a score of positive one. This scoring procedure increases the penalty for guessing incorrectly and hence, compounds women's reluctance to guess. Although the current scoring systems for programming contests do not use negative scoring, this issue is still of consequence. The scoring for the ACM International Collegiate Programming Contest (ICPC)[2] applies a penalty for an incorrect submission. This penalty can be viewed as a form of negative scoring that will increase the reluctance of women to submit a solution for which they have any doubt. While IOI style competitions do not impose any scoring penalties, it is imperative that changes to contest judging procedures do not introduce penalties for incorrect or partially correct solutions.

2.2.3. *Question Topics and Difficulty*

Educational research suggests that the phrasing of questions on examinations is often biased. For example, in her investigations on standardized testing, Rosser (1989) found that women did better on questions involving relationships, aesthetics and the humanities, while men did better on questions involving the physical sciences and business. In fact, the producer of such tests, Educational Testing Services, has acknowledged that questions must include items that are of interest to both genders, such as the humanities for women and politics or sports for men, to allow for more equal scoring, even though the topic of the item does not necessarily involve the contextual information (Dwyer, 1976). With respect to programming contests, these studies suggest that items should include

[2]See, `http://icpc.baylor.edu/icpc`, for additional detail.

contextual information that is of interest to both women and men, as is further elaborated in Section 3.2 of this paper.

Question difficulty is another potential avenue for gender differences to develop. If women were less competent at programming, one might expect that they would be able to answer only the easier questions accurately. Although we do not have direct evidence from computer science, high school mathematics testing suggests that women do not differ in competency and can perform as effectively as males. In fact, there is evidence that females answer questions more correctly than males when the questions are computationally easier but the difference disappears when the questions are more difficult (Duffy *et al.*, 1997; Hyde *et al.*, 1990).

Standardized testing, such as the SAT-M, shows that gender impacts one's ability to solve "conventional" problems (i.e., the problem is routine, textbook style, involving the application of algorithms) and "unconventional" problems (i.e., use long term math knowledge and the application of novel insight or the unusual use of familiar algorithms). Women are more likely to correctly solve conventional problems using algorithms, while men are more likely to solve unconventional problems using logical estimation and insight (Gallagher and De Lisi, 1994).

2.2.4. *Performance Anxiety*

Feedback often increases women's confidence and performance expectations. For example, Lenney (1977) found that women have lower expectations for their performance when given ambiguous feedback or no feedback. Research suggests that feedback is particularly important for women as they are more influenced by the content of feedback than are men (Helgeson, 2005).

Again, we turn to the literature on mathematics in high schools. Routinely, females tend to receive the highest evaluations in such classes (Kimball, 1989), yet perform more poorly than males on standardised tests. One explanation is that the latter environment induces stress and anxiety for some women, and consequently, leads them to perform poorly (Duffy *et al.*, 1997). This contention is supported by the idea that standardized tests often lead to the development of stereotype threat; essentially, because women are thought to perform poorly on tests involving mathematics (or, perhaps computer science), they will identify with the threat, thus verifying it and performing sub-optimally (Duffy *et al.*, 1997; Steele and Aronson, 1995).

Anxiety can be important longitudinally, especially since participants in contests are expected to repeatedly compete over time. Research on mathematics suggests that the less anxiety students feel, the higher their motivation to succeed, and consequently, the better their performance at a later point (Duffy *et al.*, 1997). As well, women experience a decrease in self-confidence when they perceive that their work will be compared to others for evaluation (Lenney, 1997). When there is no pressure (i.e., in a non-competitive environment), there is equivalent enthusiasm reported between women and men for computing activities (Astrachan, 2004). This finding is supported by data that indicates, for computing tasks, women experience decreased performance in competitive situations (Johnson *et al.*, 1985). It is likely that this performance penalty is partially the result of the increased anxiety that women experience when they are required to perform computing activities in public as opposed to in private (Robinson-Stavely and Cooper, 1990). Moreover, psychological research has indicated that women, but not men, who compete successfully against the opposite sex feel tense, nervous and anxious during the competition (Morgan and Mausner, 1973).

When it comes to motivation and participation, research in developmental psychology shows that boys often enjoy competition, which leads to higher levels of mastery of the task at hand. In contrast, girls' mastery, particularly for stereotypically non-female subjects such as math, was directly related to the degree of support they perceived from the instructor (Farmer, 1997). Thus, feedback is a particularly important tool for reducing female anxiety and increasing measurable performance.

2.3. *Strategies of Completeness*

Education research on standardized testing has shown that women use a different problem solving style than men. They are more likely to work a problem out in completion before moving on to another item, to consider more than one possible answer, and to check their answers[3]. As mentioned, women are less prone to guessing (Voyer and Saunders, 2004) and more prone to using algorithmic solution approaches (Gallagher and De Lisi, 1994). Since women desire and perform more effectively when they receive feedback (Helgeson, 2005), it is likely that women use a more thorough solution strategy as to generate personal feedback. By carefully exploring, in a systematic manner, all approaches to a problem, one can rule out ineffective approaches and thus generate some feedback in the process. As well, the likelihood that a strategy is correct increases, and the level of risk decreases, when one can filter out inappropriate responses and strategies. Thus, there are ample reasons for women to be more careful and fully explore a single problem before attempting the next.

In the domain of programming, one checks the correctness of code by performing testing. Women are therefore likely to expend more effort in testing their code in order to verify its correctness. As well, since executing a test generates feedback, there is additional incentive for women to thoroughly test contest solutions.

When not motivated by competition, one is not likely to focus on the speed in which a question is completed, which matters only for competition scoring. Instead, one may focus on the actual problem and its solution. Solving a difficult problem well can be as motivating and enjoyable as solving it quickly. Thus, as a consequence of their tendency to fully explore a problem, women are likely to find more enjoyment in solving a problem well, as opposed to solving it quickly but inelegantly.

[3] *Gender Bias in College Admissions Tests*. FairTest: The National Center for Fair & Open Testing, Cambridge, Massachusetts. `http://www.fairtest.org/facts/genderbias.htm` (last accessed 7/11/2005).

3. Mitigating Biasing Behaviours

3.1. *Feedback*

To aid in solution development, the ACM's ICPC provides a minimal feedback mechanism by providing participants with sample data. Such a mechanism is imperative to obtain female participation and could benefit by being improved. One suggestion would be to provide participants with the test data, at some points penalty, for the tests that their solution fails. To add an element of strategy, the penalty could be based on the time remaining in the contest, thus forcing participants to wait as long as possible before requesting the judging data. As well, if the data is organized by some scale of difficulty, participants might only be given the data for the first failed test case. The number of variations is endless and is certainly the subject of some interesting experimentation.

Contests that provide no correctness feedback, such as the IOI, are particularly discouraging to female participants. The addition of some feedback mechanism would certainly be an improvement. As well, the ability to see other contestants progress would decrease risk by improving the information available to make decisions about which questions to attempt, and when to abandon a question and move to another. Such information is vital to minimizing risk.

3.2. *Questioning*

Lepper and Malone (1987), in their study of computer aided learning, identified specific differences in the types of gaming activities that appealed to each gender. Not surprisingly, males preferred action and adventure games that females found to be uninteresting and unappealing. While this study examined gaming, it is equally applicable to programming contests. It is highly likely that the majority of contest questions, that are developed by male computer scientists, are more interesting to male students. As Schofield (1995) reports, girls in an advanced university level computer science class felt ridiculed and experienced increased anxiety for their inability to differentiate football and baseball teams and each sport's statistics on an assignment.

The ACM ICPC often obfuscates questions using elaborate and detailed stories. If these stories involve sports and engineering, they will often lack interest, increase anxiety, and decrease the performance of female participants. While we do not suggest the use of stereotypical "women's" issues (e.g., cooking, child-raising) as the basis for questions, there is certainly room for the use of more generic topics that are equally appealing to members of both genders.

3.3. *Practice and Motivation*

It would be unreasonable to expect an athlete to excel in competition without practice. As well, we do not expect athletes to initially practice against world-class opponents. Instead, there are developmental leagues and teams where novices can compete against opponents of an equivalent skill level. For academic-oriented competitions, there must be equivalent developmental opportunity. As women are particularly discouraged by failure in computing activities (Cooper and Weaver, 2003), it is important that their chance of success is maximized by providing them with opponents that they have a chance of defeating. While low-level competitions exist, such as the Junior-level competition of the Canadian Computing Competition[4], they are too few in number and frequency to support competitor development.

Investigations on children (e.g., Strayer and Strayer, 1976) and adults (e.g., Cashdan, 1999) have revealed that females are much less interested in hierarchical evaluation than males. Thus, for programming contests, coaches and organizers should emphasise the benefits of participation if they wish to attract females. We predict that women will be more interested in the content of a contest, including the material and skills that they will learn, and not the evaluation of their performance. Designing contests such that they build on previously learned skills will permit participants who are not as initially motivated by the competitive aspects to view the contest for its developmental opportunities and likely increase participation.

Computer programming is often viewed as a solitary activity with minimal social content. However, women are often willing to forego active or immediate socializing (e.g., watching a movie) when an event provides them with a basis for future socialization. Contests will thus be more appealing to women if they can be integrated into their social context. It is possible that, rather than encouraging the most capable girls to enter a high-school programming contest, coaches may have more success in encouraging those with a leadership role who will entice their peers to participate and thus create a group social event. Furthermore, the travel and other activities (e.g., after-contest parties) associated with programming contests are highly social in nature. If participants are unaware of these elements, they will decide on whether to participate using only stereotypical views and knowledge.

3.4. *Privacy*

While many contests require contestants to participate in open view of other contestants, it is possible that this practice is particularly discouraging to female participants. Robinson-Stavely and Cooper (Robinson-Stavely and Cooper, 1990) report that female performance significantly exceeded that of males on a computer-based task (playing the game "Zork") when in private, but was significantly exceeded by males when another person was present. Of course, it is not realistic to give each contest participant their own, private, facility for computing competitions. However, the magnitude of the performance decrease is highly distressing. Contest organizers should consider this effect and seek to afford female participants a preferred seating location. It is likely that even simple interventions, such as giving females the more private seats at the back of the room, will have an impact on their performance.

Females also benefit from working in single-sex groups (all female environment) (Light *et al.*, 2000) when performing computing tasks. Again, when possible, it would

[4]See, `http://contest-cemc.uwaterloo.ca/ccc`, for additional detail.

benefit female participants to be located together, preferably in their own room, during competitions. Of course, this is not always possible, but for contest organizers to be aware of this affect and to attempt beneficial interventions is an important step towards increasing female participation.

Privacy can also be increased using anonymity. For example, when providing a real-time scoreboard during a contest, participants can be identified by a randomly assigned identifier that only the participant knows. Each contestant can identify their standing against the others, but is secure in the knowledge that the others cannot identify their position. When the winners are announced, those contestants that did poorly are spared from embarrassment.

3.5. *Teams and Individuals*

Johnson *et al*. (1985) examined student learning and test performance after engaging in computer-based learning alone and in cooperative groups. While, as previously mentioned, female performance is negatively impacted by the presence of others, the impact is decreased when females are in a cooperative environment. Working alone and being solely responsible for one's results increased anxiety and reduced performance, as compared to working with others and sharing responsibility for results. The impact of the finding is clear for programming contests. Female participants will prefer team-based cooperative contests over individual contests. For cultures that prize cooperation over competition (Galliano, 2003) this finding is also critical. When it is not possible for team-based competitions, anonymising the participants may help alleviate some of the anxiety experienced by contestants.

Computer science suffers from a lack of female role models (Jepson and Perl, 2002). This situation is certainly reflected in programming competitions where the majority of competitors, coaches and judges are male. It has been reported that at some ACM ICPC events, the only female participants have been from all-female colleges (e.g., Smith and Wellesley Colleges). While it is known that women function better in a single sex environment (Cooper and Weaver, 2003), students from these schools will also have an advantage in that older students can serve as role models for younger students on contest teams. To encourage the development of role models, the participation of women should be recognized. In some contests, the best-placed team from a particular geographic region is identified. Perhaps this practice should be extended and the best placed female competitors also recognized. There are certainly pitfalls to such a practice, but potential solutions to this issue warrant future investigation.

3.6. *Scoring*

Most scoring systems are particularly biased in favour of male participants. As mentioned, females are more risk averse and prefer to receive some level of feedback about their performance. The lack of feedback in most current systems, such as not reporting the number of test cases that a solution passed or failed, prevents an incremental approach that permits contestants to submit partial solutions and then refine the solution. Furthermore, it is known that women tend to prefer a bottom-up program development strategy (Turkle and Papert, 1990) that uses incremental development, likely due to the reduced risk of such a strategy. One need not make the feedback system trivial and could, for example, use an approach based on the strategy game "Mastermind" to report correctness.

For many contests, the primary judging criteria, after correctness, is the time taken to achieve a correct solution. For female participants, it has been shown that time pressure will negatively affect performance. We are not suggesting that time is not a significant indicator of programming skill. It is clear that a correct solution developed in less time is, in some ways, better than one that takes longer to develop. However, the quality of the developed code, the algorithmic elegance employed, and the robustness with respect to invalid data are also indicators of a solution's quality. In general, our ideas can be summarized by the phrase "good code should be rewarded." That is, sound practices and their products should be rewarded and, for equally correct solutions, participants that produce high quality solutions should not be excessively penalized because of the extra time that this quality required.

The need for speed that current scoring systems use often forces contestants to abandon programming practices that they have been taught are desirable. For many students, this change causes contests to become artificial and less attractive. As previously discussed, women tend towards using algorithmic approaches to problem solving that are based on the familiar. Practices that vary from those taught in the classroom are likely to have a more negative impact on women than on men. Thus, we suggest that code quality is also an attribute worth rating. Students are taught to select meaningful variable names, format code appropriately and so on. We have seen these practices abandoned in contest environments as they can hinder the speed of program development. Women are less likely to change their programming style for an unfamiliar setting, and are thus harmed further by a scoring system based primarily on time. The points awarded for code quality need not be many, but their inclusion in a more encompassing scoring system is something that should be considered.

In our experience with ACM ICPC teams, we have observed that female participants often attempt to develop elegant (i.e., efficient) solutions to problems rather than rapidly obtaining a working solution. As we have shown, the available literature supports this observation. While the "time to market" is an important element of real-world software development, code that is slow and inefficient will soon drive users to seek other products. Programming contests often impose execution constraints to prevent the submission of grossly inefficient solutions. However, a binary judgment of efficiency is perhaps much too coarse a measure. As well as development time, execution time could also be recorded and used to determine overall scoring.

In programming contests, it is typical to instruct participants to expect the judging data to be correctly formed. However, as any experienced coach knows, errors in judging data are common-place and an occurrence for which one must be prepared. Students who develop solutions that detect these errors should be rewarded. In programming classes, robustness is taught and good students develop code that detects invalid input. It does not

infon v.1998/05/11 Prn:17/03/2006; 14:07 F:INFE073.tex; VTEX/RRR p. 13

seem desirable to suddenly suggest that these good habits are no longer appropriate and are not worthy of reward. Risk averse students, and in particular female students, should not be subject to time penalties when they continue to employ the good programming practices they have been taught. While the rewards (points) for robustness should likely be minimal, it should be valued and earn the competitor some credit.

In an improved contest scoring system, the combination and weight that should be attributed to each of these issues is a topic for future discussion. We advocate the development of a more inclusive system that rewards good programming practices, as well as the speed of program development. Even if such a system does not turn out to be feasible, it is still possible to reward competitors for alternative solutions. For example, a simple improvement would be the rewarding of subsidiary prizes for "Most Efficient Solution" and "Best Quality Solution". Such ideas are not completely novel, as the Programming Challenges web site[5] reports the run-time for each submission and the best run-time for each question.

Although, as a mathematical discipline, computer science tends towards "absolute" judging, there are merits to a "relative" approach. As suggested by Colwell (2005), "both relative and absolute judging are essential to achieving the right balance of fun and education (in science fairs)". As an example, he suggests the case where all contestants provide unfeasible solutions. Rather than give a score of 0 to everyone, it is possible to relatively rate the submissions on their merits. For lower-level competitions that seek to foster involvement through making the event fun and enjoyable, relative judging can provide the motivation for continued involvement, rather than the disincentives of a 0 score.

The IOI uses "relative judging" for some questions and assigns the best solution full marks, with other submissions graded relative to the best solution. This scoring approach can be beneficial to women if scores are based on program performance and not simple development time. An alternative form of relative judging is to evaluate submissions relative to a predetermined standard. For example, submissions could be evaluated against a best known, or theoretically best, solution. Women may find this variation more attractive if they focus on equaling the standard as opposed to crafting a better solution than other competitors. Research is needed to determine the least biased and most equitable form of relative judging.

4. Conclusion

Over a decade ago at the 1991 IOI in Greece, it was suggested that, "Girls should be strongly encouraged to participate".[6] In 1992, a Dutch proposition for all teams to have at least one female team member was not endorsed. While the proposal was well intended, we believe that it would likely have failed to achieve its goal. For women to succeed at the IOI, they must enter the competition with both the competence and confidence to win. For a world-class event, participants must have world-class skills; selection based on gender

[5] See `http://www.programming-challenges.com` for additional detail.
[6] See, `http://olympiads.win.tue.nl/ioi/ioi92/report.html`, for additional detail.

would not guarantee that female participants had the necessary skills and would thus set them up for failure, exacerbating and not solving the initial problem.

The available literature makes the focus of this paper clear: Our current practices for programming contests do little to encourage participation by women, and presumably other demographic groups. However, simply changing a contest's management, scoring or environment is not going to immediately increase womens' participation or address current social and cultural issues. Complex social problems have complex solutions with many elements and that require considerable time to enact. As part of the solution, contest organizers must slowly and carefully identify areas where improvement in gender equity can be made. Competitive computing, by presenting an unbiased and fair environment, has the potential to be a leader in addressing gender inequity in computing disciplines. We suggest that it is time to begin exploring the changes that are needed.

Although the culture and organization of existing contests is well established, there is certainly room for improvement. In this paper, we suggest several interventions that may aid in improving female participation and improve the ability of contests to serve as a recruiting mechanism for higher-level, post-secondary academic institutions. These suggestions, while based on sound theory, should now be examined in actual contests to determine their effectiveness and to identify other areas where improvements can be made. Contests provide an important venue for the best and brightest to distinguish themselves; it should be our goal to ensure that exclusionary barriers are minimal and that all of our best are equally capable of earning the accolades they deserve.

References

Astrachan, O. (2004). Non-competitive programming contest problems as a basis for just-in-time teaching. In *Proc. ASEE/IEEE Frontiers in Education Conference*. Savannah, Georgia, pp. T3H/20–T3H/24.

Byrnes, J., and S. Takahira (1993). Explaining gender differences on SAT-math items. *Developmental Psychology*, **29**, 805–810.

Cashdan, E. (1999). Are men more competitive than women? *British Journal of Social Psychology*, **37**, 213–229.

Colwell, B. (2005). Judging science fairs. *IEEE Computer*, **38**(7), 12–15.

Cooper, J., and K. Weaver (2003). *Gender and Computers, Understanding the Digital Divide*. Lawrence Erlbaum and Associates, Mahwah, New Jersey.

Daly, M., and M. Wilson (2001). Risk-taking, intrasexual competition, and homicide. *Nebraska Symposium on Motivation*, **47**, 1–36.

Duffy, J., G. Gunther and L. Walters (1997). Gender and mathematical problem solving. *Sex Roles*, **37**, 477–494.

Dwyer, C. (1976). Test content and sex differences in reading. *The Reading Teacher*, 20–24.

Farmer, H. (1997). Women's motivation related to mastery, career salience, and career aspiration: A multivariate model focusing on the effects of sex role socialization. *Journal of Career Assessment*, **5**, 355–381.

Fennema, E., and P. Peterson (1987). Effective teaching for girls and boys. In D. Berliner and B. Rosenshine (Eds.), *Talks to Teachers*. Random House, New York, New York.

Gallagher, A., R. De Lisi, P. Holst, A. McGillicuddy-De Lisi, M. Morely and C. Cahalan (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology*, **75**, 165–190.

Gallagher, A., and R. De Lisi (1994). Gender differences in scholastic aptitude test – mathematics problem solving among high-ability students. *Journal of Educational Psychology*, **86**, 204–211.

Galliano, G. (2003). *Gender, Crossing Boundaries*. Thomson Wadsworth, Belmont, California.

Goldstein, D., D. Haldane and C. Mitchell (1990). Sex differences in visual-spatial ability: The role of performance factors. *Memory and Cognition*, **18**, 546–550.

Helgeson, V. (2005). *Psychology of Gender* (2nd edition). Prentice-Hall, Upper Saddle River, New Jersey.

Hoyenga, K., and K. Hoyenga (1993). *Gender-Related Differences*: *Origins and Outcomes*. Allyn and Bacon, Needham Heights, Massachusetts.

Hrdy, S. (1999). *The Woman That Never Evolved*. Harvard University Press, Cambridge, Massachusetts.

Hughes, L. (1988). "But that's not really mean": Competing in a cooperative mode. *Sex Roles*, **9**, 669–687.

Hyde, J., E. Fennema and S. Lamon (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, **107**, 139–155.

Jepson, A., and T. Perl (2002). Priming the Pipeline. *ACM SIGCSE Bulletin "Inroads"*, **34**(2), 36–39.

Johnston, M., and C. Crawford (1999). Stigmatising women's aggressive behaviour: Who does it benefit and why? *Behavioral and Brain Sciences*, **22**, 226–227.

Johnson, R., D. Johnson and M. Stanne (1985). Effects of cooperative, competitive, and individualistic goal structures on computer-aided instruction. *Journal of Educational Psychology*, **77**, 668–677.

Kimball, M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, **105**, 198–214.

Lenney, E. (1977). Women's self-confidence in achievement setting. *Psychological Bulletin*, **84**, 1–13.

Lepper, M., and T. Malone (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. Snow and M. Farr (Eds.), *Aptitude, Learning and Instruction: Vol. 3. Conative and Affective Process Analysis*. Lawrence Erlbaum and Associates, Mahwah, New Jersey, pp. 255–286.

Light, P., K. Littleton, S. Bale, R. Joiner and D. Messer (2000). Gender and social comparison effects in computer-based problem solving. *Learning and Instruction*, **10**, 483–496.

Morgan, S., and B. Mausner (1973). Behavioral and fantasized indicators of avoidance of success in men and women. *Journal of Personality*, **41**, 457–470.

Robinson-Stavely, K., and J. Cooper (1990). Mere presence, gender and reactions to computers: Studying human-computer interaction in the social context. *Journal of Experimental Social Psychology*, **26**, 168–183.

Rosser, P. (1989). *The SAT Gender Gap: Identifying the Causes*. The Center for Women Policy Studies, Washington, D.C.

Schofield, J. (1995). *Computers and Classroom Culture*. Cambridge University Press, Cambridge, Massachusetts.

Steele, C., and J. Aronson (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, **69**, 797–811.

Stockard, J., J. van de Kragt and P. Dodge (1988). Gender roles and social dilemmas: Are there sex differences in cooperation and in its justification? *Social Psychology Quarterly*, **51**, 154–163.

Strayer, F., and J. Strayer (1976). An ethological analysis of social agonism and dominance relations among preschool children. *Child Development*, **47**, 980–989.

Turkle, S., and S. Papert (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs*, **16**, 128–157.

Vandenburg, S., and A. Kuse (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perception and Motor Skills*, **47**, 599–604.

Voyer, D., and K. Saunders (2004). Gender differences on the mental rotations test: A factor analysis. *Acta Psychologica*, **117**, 79–94.

M. Fisher received her PhD in psychology from York University in 2004 and is currently an assistant professor in the Department of Psychology at Saint Mary's University. She lectures on the psychology of gender and evolutionary psychology. Her research interests include female intrasexual competition, social relationships, and gender differences in cognition.

A. Cox received his PhD in computer science from the University of Waterloo in 2002. He is currently an assistant professor in the Faculty of Computer Science at Dalhousie University. He is also a member of the problem committee for the Canadian Computing Contest and has coached Dalhousie's ACM ICPC teams. His research focuses on the human factors of computing, program comprehension, and software tool development and evaluation.

Lyčių lygybė ir programavimo varžybos: diskriminacijos mažinimas

Maryanne FISHER, Anthony COX

Individai vienas nuo kito skiriasi daugeliu dimensijų, susijusių su lyties, asmenybės tipo ar kultūrinių skirtumų aspektais. Tokiu būdu galima teigti, jog programavimo varžybos, pasižyminčios ribota struktūra (t.y. vertinimo taškais sistema, klausimų stiliumi) yra labiausiai mėgstamos bei patraukia specifinį individų, kurių savybės geriausiai atitinka tokią struktūrą, ratą. Straipsnyje keliama prielaida, jog paslankesnė varžybų struktūra, būdama mažiau šališka atskirų asmens savybių atžvilgiu, leistų joms tapti labiau patraukliomis platesnei auditorijai. Be to, pritaikius varžybas platesnei auditorijai, jos galėtų labiau pasitarnauti skatindamos didesnį mokinių ratą pasirinkti studijuoti informatikos disciplinas. Straipsnyje koncentruojamasi ties su lytimi susijusiais skirtumais, taip pat tiriamas varžybų struktūros poveikis merginoms.