ABSTRACT: Nearly 79,000 questions sent to an Internet-based Ask-A-Scientist site during the last decade were analyzed according to the surfer’s age, gender, country of origin, and the year the question was sent. The sample demonstrated a surprising dominance of female contributions among K-12 students (although this dominance did not carry over to the full sample), where offline situations are commonly characterized by males’ greater interest in science. This female enthusiasm was observed in different countries, and had no correlation to the level of gender equity in those countries. This suggests that the
Internet as a free-choice science-learning environment plays a potentially empowering and democratic role that is especially relevant to populations that are traditionally deprived of equal opportunities in learning formal science. However, worldwide, girls’ interest in submitting questions to scientists dropped as they grew older relative to the boys’ interest, and the stereotypically gendered science interests persisted in this environment as well. The strengths and limitations of using free-choice Web-based data sources for studying youth interest in science are discussed. © 2008 Wiley Periodicals, Inc. Sci Ed 1–30, 2008

INTRODUCTION

An earlier paper in this journal (Baram-Tsabari, Sethi, Bry, & Yarden, 2006) drew upon the questions sent to an Ask-A-Scientist site during a 3-month period at the end of 2004 to identify the scientific interests of children. The present paper draws upon the same source to analyze nearly 79,000 questions gathered since the establishment of the Internet site in the mid-1990s (www.madsci.org), to learn about the interactions between age, gender, country of origin, and interest in scientific topics. More specifically, we wanted to determine how the participation of female contributors changed between countries, with age and with time, and how the scientific interests, reflected by the subject of the questions, changed with age and time, between genders, and between countries.

THEORETICAL FRAMEWORK

Students’ Questions

Although question asking is a basic requirement for the performance of scientific research and meaningful learning, the way in which science lessons are usually conducted does not stimulate question asking by students, and questions are posed mainly by the teachers (Allison & Shrigley, 1986; Dillon, 1988; Dori & Herscovitz, 1999; Graesser, Person, & Huber, 1992; Marbach-Ad & Sokolove, 2000). Requests for meaningful explanations are relatively infrequent in K-12 classrooms, and students at all grade levels (K-12) generally ask the same number of questions (Good, Slavings, Harel, & Emerson, 1987)—approximately 1% of the questions asked in class (Graesser et al., 1992).

Many researchers have recommended strategies to encourage question asking by students, such as one-on-one tutoring sessions (Graesser & Person, 1994), discussion (King, 1994), cognitive conflict (Allison & Shrigley, 1986), real-world problem-solving activities (Chin, Brown, & Bruce, 2002; Zoller, 1987), case studies (Dori & Herscovitz, 1999), biotechnology-focused modules (Olsher & Dreyfus, 1999), use of written questions (Pedrosa de Jesus, Teixeira-Dias, & Watts, 2003), and learning using adapted primary literature (Brill & Yarden, 2003).

In spite of the effort to encourage question asking by students, students are more often expected to answer questions than to ask them in the typical classroom setting (Chin, 2004), and the common situation in science classes is still the one described by Dillon (1988): “Children qua (=as) students do not ask questions. They may be raising questions in their own mind . . . but they do not ask questions aloud in the classroom.” Researchers attribute this situation to a classroom atmosphere in which revealing a misunderstanding may render the student vulnerable, open to embarrassment, censure or ridicule (Pedrosa de Jesus et al., 2003). Students have described their teachers’ responses to their questions as “put-offish” or even annoyed, and their classmates’ reaction as “intolerant” (Rop, 2003).

The overall paucity of student question asking has resulted in relatively few studies of pupils’ questions, simply because researchers have not been able to find enough of them to examine (Maskill & Pedrosa de Jesus, 1997; Pedrosa de Jesus et al., 2003; Watts &
Alsop, 1995). Good et al. (1987) specifically note the absence of comparative studies of student-generated questions across different grade levels using the same methodology.

Dim as the picture of student-generated questions may be, research, as well as life experience, tell us that students are capable of asking many questions when given the opportunity (Costa, Caldeira, Gallastegui, & Otero, 2000). Therefore, a better way to study children’s questions might be to look for them where they are being asked fluently and voluntarily. Learners ask questions when they feel secure (Watts, Gould, & Alsop, 1997), and one place offering such security is free-choice science-learning settings on the Internet.

The Web as a Free-Choice Science-Learning Environment

Examining free-choice science-learning environments can provide knowledge about the natural setting in which people learn in a self-directed, self-motivated, voluntary way, guided by individual needs and interests (Falk & Dierking, 2002). An example of such a free-choice setting is the World Wide Web, which is the primary source for news and information about science for 20% of Americans, second only to television (41%). Moreover, if people need information on a specific scientific topic, the Internet is the primary source to which they turn (Horrigan, 2006). Two thirds of Internet users say they have come upon news and information about science when going online for other purposes, and half of all Internet users have been to a Web site that specializes in scientific content (Horrigan, 2006).

While access to the Internet grows exponentially, American students are already wired: in fall 2005, nearly 100% of public schools in the United States had access to the Internet, and 94% of the instructional rooms in those schools had Internet access (National Center for Education Statistics [NCES], 2006). In 2005, 87% of all youth between the ages of 12 and 17 had used the Internet, and 68% of all teenagers had used the Internet at school (Rainie & Hitlin, 2005). In general, youth hold positive attitudes and exhibit self-confidence with respect to Internet use (Fidel et al., 1999; Watson, 2004). Their use of the Web ranges from researching for school assignments to communicating with others and exploring their personal interests (Baram-Tsabari et al., 2006; Bilal, 2004; Hirsh, 1999; Levin, Arafah, Lenhart, & Rainie, 2002; MaKinster, Beghetto, & Plucker, 2002; Weigold & Treise, 2004).

Science Web sites may influence young people’s life-long interest in science and their appreciation of its beauty and importance (Weigold & Treise, 2004). However, a listing of the 15 sites most visited by teenagers did not include any sites related to science, technology, or even education in general. Weigold and Treise concluded that teenagers usually go to the Web to have fun and interact with others, but only occasionally use the Web to learn.

There are mixed findings regarding the role of gender in using the Web as a free-choice environment for science learning. Although boys have more formal and out-of-school experience using computers and the World Wide Web (Kafai & Sutton, 1999; Shashaani, 1994), more girls prefer this type of lesson over traditional classroom-based science learning (Leong & Al-Hawamdeh, 1999). Ching, Kafai, and Marshall (2000) found that configuration of social, physical, and cognitive gender-equitable spaces contributes to a positive change in girls’ level of access to programming activities. The American Association of University Women (AAUW; 2004) describes girls and women as being attracted to the communicative aspects of online interactions, and therefore recommends online projects as a means of promoting gender-equitable participation.

A former study that used questions from an Ask-A-Scientist source found that among 4th–12th graders, girls asked most of the questions in contributions arriving from the United States, Canada, and the UK, but not from the other countries surveyed (Baram-Tsabari et al., 2006). This female majority contradicts previous female-to-male ratios obtained from a scientific Internet site based in Italy (Falchetti, Caravita, & Sperduti, 2003), a UK-based
science line (K. Mathieson, personal communication, April 2, 2004), and science and technology questions at an Israeli children’s Web site (Baram-Tsabari & Yarden, 2005). At the interactive Web site Whyville, which was designed to engage students in socially interactive, entertaining, and educational activities that include inquiry science, most users were found to be girls (73% of the regular Whyville users who answered a survey), contrary to what might be expected from a science-oriented program (Aschbacher, 2003). The question of female usage of the Web as a free-choice environment for science learning should be viewed in the context of females’ general reluctance to use media that foster informal learning about science (National Science Foundation [NSF], 2004; Nisbet et al., 2002), and to take part in extracurricular science experiences (Greenfield, 1998).

Research on children’s use of the World Wide Web for learning has generally been conducted in school settings (e.g., Bilal, 2001; Fidel et al., 1999; Guinee, 2004; Rogers & Swan, 2004; Slotta, 2004). The Web is seen by educators as a site for student inquiry in science, which allows students to pursue questions of personal interest (McCready Wallace, Kupperman, Krajcik, & Soloway, 2000), since an effective search is also an exercise in inquiry and critical thinking (Brem & Boyes, 2000). However, most students have difficulty formulating and modifying search queries (Bilal, 2004; Hirsh, 1999; MaKinster et al., 2002; Wallace, Kupperman, Krajcik, & Soloway, 2000; Watson, 2004), and many of them fail to construct an accurate and broad understanding following an online inquiry (Hoffman & Krajcik, 1999).

Furthermore, children do not tend to question the accuracy of the information they find on the Web (Hirsh, 1999; Russell, Weems, Brem, & Leonard, 2001; Schacter, Chung, & Dorr, 1998; Wallace et al., 2000). Such skepticism is sorely needed, as Keating, MaKinster, Mills, and Nowak (1999) found that as few as 30% of the search results for science concepts actually contain at least a short operational definition or graphic display of the concept, and many of the sites contain misconceptions. Another major problem is that students believe that they should be able to find answers to complex questions on specific Web pages, instead of researching to form an answer (Soloway & Wallace, 1997). To sum up, although the Internet has the potential to greatly facilitate positive changes in education, its use in school is sporadic, peripheral to the core curriculum, and simple and obvious in nature (Schofield, 2005).

Consequently, students report that there is a substantial disconnectedness between how they use the Internet for school and how they use it during the school day and under the teacher’s direction. For the most part, students’ educational use of the Internet occurs outside of the school day, outside of the school building, away from teacher direction (Levin et al., 2002). Steinkuehler (in press) suggests that to understand the current and potential capacities of technology for cognition, learning, literacy, and education, we must look at contexts outside the current formal educational system rather than those within. The reason for this is that what students do with online technologies outside the classroom is not only markedly different from what they do with them in schools but it is also more goal-driven, complex, sophisticated, and engaged. Hence, it might prove more fruitful to study children’s use of the World Wide Web for learning in free-choice settings, rather than in school settings.

When children are using the Internet to research their interests, some of their complex questions are better answered by experts than by a list of directories or sites. This type of service is offered on the Web by human-mediated question-and-answer sites and they are sometimes referred to as “expert services” (Janes, Hill, & Rolfe, 2001) or “Ask-A” services, such as “Ask-A-Scientist” (Lankes, 1999). Such sites usually maintain searchable public archives in which previously answered questions are returned as search results, thus making this archive a resource for their users (Pomerantz, Nicholson, Belanger, & Lankes, 2004). This study uses the archives of such a Web site as a data source for identifying interests in science.
Interest in Science

Adolescents’ decisions about the contents and directions of their educational training are influenced to a high degree by the topic-related interests they developed in the preceding years (Krapp, 2000). Interest was the primary reason for choosing to enroll in an advanced science class among Israeli (Levy, 2003), Swedish (Lindahl, 2007), American and Australian students (Kahle, Parker, Rennie, & Riley, 1993). Interest does not only affect the choice of courses and career but also the ability to learn. Research indicates positive relationships between interest and a wide range of learning indicators (Pintrich & Schunk, 2002; Schiefele, 1998).

Regardless of the importance of interest, the current situation in science education was summarized by a Swedish student in the following manner: “The trouble with school science is that it provides uninteresting answers to questions we have never asked” (Osborne, 2006). The untested assumption is that the more we know about students’ interests, enthusiasms, dislikes, beliefs, and attitudes, the more feasible it will be to develop school science curricula that will engage their attention and help reduce long-standing gender and other differentials (Jenkins, 2006).

Within free-choice science education, some attention is paid to the audience’s interests. Front-end studies are regularly used by museums to initiate a dialogue with visitors, enabling the exhibition developers to learn about the audience’s assumptions, understandings, attitudes, beliefs, and interests (Dierking & Pollock, 1998). Another free-choice institute that may benefit from tuning in to its audience is the public library. Cooper (2004) studied children’s choices for inclusion in a hypothetical library, based on the understanding that children’s information interests and needs differ from those of adults, even those adults who have the best intentions when it comes to understanding and predicting children’s interests.

When high school students are asked to indicate their interest in learning about various topics in their science classes, they choose topics such as disease (cancer and HIV/AIDS), drugs (therapeutic and recreational), biological and chemical weapons, the ozone layer, and greenhouse gases. Yet the usual high school science curriculum does not address these topics (Kwiek, Halpin, Reiter, Hoeffler, & Schwartz-Bloom, 2007). Overall, biology is the most popular science subject among students (Baram-Tsabari et al., 2006; Baram-Tsabari & Yarden, 2005; Dawson, 2000; Murray & Reiss, 2005; Osborne & Collins, 2000; Qualter, 1993), especially among females (see next section for gender aspects of interest in science). Within the field of biology, students display significant changes in the structure of their interests with age: human biology becomes important while interest in plants and animals decreases (Baram-Tsabari et al., 2006; Baram-Tsabari & Yarden, 2005, 2007; Osborne & Collins, 2000; Stawinski, 1984; Tamir & Gardner, 1989).

The profiles of students’ experiences and interests vary strongly between countries (Sjøberg, 2000). Results from the ongoing international “Relevance of Science Education” [ROSE] project indicate that similarities among countries in students’ responses, regarding what science topics they would like to learn, are first determined by geographical proximity, and next by the level of development, indicated by the UN’s human development index (Sjøberg & Schreiner, 2005). However, these differences might be much smaller among Western countries. Lyons (2006), for example, found a remarkable similarity in the experiences of school science reported by high school students in Sweden, England, and Australia. This analysis revealed that students in different educational and national contexts were not only experiencing very similar high school science classes but also identifying similar problems and responding in similar ways.
Gender and Interest in Science

To successfully address the needs and interests of underrepresented groups, we need to know not only what works, but what works for whom (AAUW, 2004). Research has provided insight into these issues, especially on the role of gender in predicting scientific interest. Most of these insights are based on data collected in school science settings.

The wealth of data regarding boys’ and girls’ interest in science can be summarized in the following manner: boys in general are more interested in science than are girls (Gardner, 1975, 1998; Miller, Slawinski Blessing, & Schwartz, 2006), especially in the fields of physics and technology. Girls, on the other hand, are more interested in biology than boys. Chemistry is equally interesting to both genders. These findings (or parts of them) have been repeated in many countries, including Scotland (Stark & Gray, 1999), Australia (Dawson, 2000; Kahle et al., 1993; Woodward & Woodward, 1998), the United States (Farenga & Joyce, 1999; Jones, Howe, & Rua, 2000), England (Murphy & Whitelegg, 2006; Osborne & Collins, 2001; Spall, Barrett, Stanisstreet, Dickson, & Boyes, 2003), Israel (Friedler & Tamir, 1990; Trumper, 2006) and Germany (Hoffmann, 2002), and in international studies such as “Science and Scientists” (SAS) (Sjøberg, 2000) and ROSE (Sjøberg & Schreiner, 2002).

The ROSE studies conducted in Denmark (Busch, 2005), England (Jenkins & Nelson, 2005), Norway (Schreiner, 2006), and Finland (Lavonen, Juuti, Uitto, Meisalo, & Byman, 2005) found that girls’ interests were focused on health, medicine, the body, the mind, and well being, whereas boys wished to learn more about the dramatic aspects of physics and chemistry, and how technology works. This gender gap in interest is also apparent among female students who are interested in science, as can be inferred from the polarized enrollment in elective biology and physics courses (Murphy & Whitelegg, 2006; Zohar, 2003) within the science-attentive student body. It is also evident in science-interest studies that use senior high school science students as a sample. For example, Osborne and Collins (2001) surveyed students’ views on school science using focus groups of 11th graders who intended to continue with their science studies, and those who did not. Girls in both groups made many more negative comments about physics than did boys. Thus, it seems that the increasing access of female students to the traditionally masculine science subjects is being accompanied by the emergence of biology as a feminine niche in science (Ayalon, 1995).

Some researchers have suggested that the basis of these stereotypically gendered interests is an inborn trait which hard-wires the average girl for empathy, while the average boy is predominantly hard-wired for understanding and building systems (Baron-Cohen, 2003). Other studies, however, did not find any such difference (Hyde & Linn, 2006). A landmark MRI study of normal brain development (Waber et al., 2007) found that mental performance differs very little with gender. In her review, Spelke (2005) states that

Thousands of studies of human infants, conducted over three decades, provide no evidence for a male advantage in perceiving, learning, or reasoning about objects, their motions, and their mechanical interactions. Instead, male and female infants perceive and learn about objects in highly convergent ways. (p. 952)

Additional explanations, which do not assume an inborn gender difference, were subsequently suggested to explain girls’ lack of interest and underrepresentation in science; these are traditionally divided into the three, somewhat overlapping categories suggested by Kelly (1978):

1. Cultural explanations, which may be referred to as “socialization explanations,” include the masculine image of science, which is seen years prior to the actual
2. *Attitudinal* explanations refer to girls’ negative attitudes toward science and pursuing a science-related career (Crettaz von Roten, 2004; Kahle & Lakes, 1983; Kelly, 1978; Miller et al., 2006; Simpson & Oliver, 1985; Weinburgh, 1995). Some factors, such as misuse, difficulty and masculine image, which were brought up to explain girls’ less favorable attitudes toward science, apply more strongly to the physical sciences (Kelly, 1978).

3. *Educational* explanations include school-related parameters, such as enrollment and achievement in mathematics classes, class atmosphere, teaching and assessment methods traditionally used in physics classes (Zohar & Bronshtein, 2005), gender-related differences in the notion of what it means to understand physics (Stadler, Duit, & Benke, 2000; Zohar, 2003), and science curricula that are heavily biased toward the interests, knowledge, and abilities of boys (Hoffmann, 2002; Nair & Majetich, 1995). Haussler, Hoffmann, Langeheine, Rost, and Sievers (1998), for example, identified five domains of interest in physics; only one of them—physics as a scientific enterprise for its own sake—is overwhelmingly dominant in physics classrooms. Other domains, such as how science can serve humankind and explanations of natural phenomena, which are of more interest to girls, are almost nonexistent (Haussler et al., 1998). Despite international reports of educational success for girls, very little has in fact changed over the past few decades with respect to their science and mathematics subject choice (van Langen, Rekers-Mombarg, & Dekkers, 2006). For a comprehensive review of recent research on girls’ participation in school physics, see Murphy and Whitelegg (2006).

Evidence from free-choice science-learning settings indicates that the polar pattern of girls’ relatively high interest in biology and boys’ relatively high interest in physics is similar to the situation described within formal science education. The gender gap is already evident among young elementary school children, before biology and physics have been identified as such, and it persists all the way into adulthood (Baram-Tsabari & Yarden, 2008).

The gender effect on science-related attitudes and beliefs is not homogeneous across measures, science-content areas, racial or socioeconomic groups (Kahle et al., 1993), or cultural or situational contexts (Linn & Hyde, 1989). However, stereotypical male and female interests seem to cross borders and cultures. The SAS project, for example, found strong similarities between the lists of Norwegian and Japanese science topics favored by boys and girls, despite the strong cultural differences between these two countries (Sjøberg, 2000). Similar findings were obtained from a comparison between Israeli and international children’s spontaneous interests (Baram-Tsabari et al., 2006; Baram-Tsabari & Yarden, 2005).

This similarity is also valid for enrollment rates of women in science- and technology-related occupations. In Egypt, for example, a survey by the Supreme Council of Universities for 1995–1996 reports that in disciplines such as pharmacy and dentistry, more than 40% of the faculty are women; in the sciences, 25% of the faculty are women, but this decreases to less than 10% in the engineering and technology departments. These statistics are very
similar to those for U.S. universities, where women constitute 50% of the health sciences faculty, 23.8% of the biological sciences faculty, and 6.1% of the engineering faculty (Hassan, 2000). In most OECD countries, the proportion of women choosing advanced science and technology studies remains below 40%, and the choice of discipline is highly gender dependent (Organisation for Economic Co-operation and Development [OECD], 2006).

Contrary to expectation, gender differences are not smaller in technologically advanced countries, which foster mass education and equity legislation, or in advantaged socioeconomic groups (Steinkamp & Maehr, 1984). To list a few recent examples, in Latin America and the Caribbean, women account for 46% of the reported number of researchers, while their share falls to 15% in Asia and about 30% in Africa. In Europe, 32% of the researchers are women, with only five countries reaching gender parity (UNESCO Institute for Statistics, 2006). Among school students, gender differences in science achievements are higher for fourth and eighth grade students from the Netherlands, compared to students from Cyprus and Latvia (Martin, Mullis, Gonzalez, & Chrostowski, 2004; UNESCO Institute for Statistics, 2005).

Females, more than their male peers, tend to lose interest in science as they grow older, mainly during the middle school and high school years (George, 2006; Greenfield, 1998). American girls’ attitude to science was found to become increasingly negative with age (Kahle & Lakes, 1983), a finding that was repeated among Israeli students (Friedler & Tamir, 1990; Shemesh, 1990). Furthermore, a significant decrease in the number of American girls submitting science questions to an Ask-A-Scientist site occurred during the transition from junior to senior high school (Baram-Tsabari et al., 2006). A study conducted in Germany also found a difference in the way interest in physics develops with age: girls, but not boys, find physics as a school subject less and less interesting as they grow older (Hoffmann, 2002; Hoffmann & Haussler, 1998).

**RESEARCH APPROACH**

Despite recent efforts to harmonize statistical information on education at the international level, current datasets do not allow for a full comparative analysis, and may lead to conflicting interpretations (OECD, 2006). It is rather difficult to use the existing data to examine interactions between gender, age, country of origin, and interest in science, as they were not specifically and deliberately addressed in advance by the researchers. The international studies SAS and ROSE describe interactions between country and gender, but they do so for a single age group (15-year-old students) in a formal science setting using questionnaires.

Interest in science has traditionally been identified using written questionnaires that rely on adultcentric views of what subjects should be meaningful to students. Cook-Sather (2002) advocates the notion that there is something fundamentally amiss about building and rebuilding the educational system without consulting those it is ostensibly designed to serve. We believe that relying on children’s spontaneous ideas and questions will enable more rapid progress toward the incorporation of their views than using their responses to an adult-written questionnaire. Therefore, we recently suggested using children’s self-generated science-related questions as a tool to probe students’ scientific interests (Baram-Tsabari et al., 2006; Baram-Tsabari & Yarden, 2005, 2007, 2008). Self-generated questions can help reveal the asker’s reasoning, alternative views, and interests (Biddulph, Symington, & Osborne, 1986), and studying students’ questions can make teachers aware of what students are interested in and what they want to know about a given topic (Chin & Chia, 2004).
Web-Based Research

The potential for online experimental laboratories for the social and behavioral sciences was already described a decade ago, in a 1997 NetLab report (NSF, 1997) which suggested that the Web could enable experiments to (i) be scaled up to include hundreds or even thousands of subjects; (ii) cross many boundaries, bringing new population samples into the laboratory; (iii) mimic lengthy time periods in which subjects interact with one another over long intervals; and (iv) make laboratory experimentation a part of the routine education of undergraduates. Ten years later, in a *Science* paper entitled “The scientific research potential of virtual worlds,” Bainbridge (2007) introduced a number of possible research methodologies for Web-based research, including formal experimentation, observational ethnography, and quantitative analysis.

The actualization of the potential of Web-based research has been rather modest to date: the overall incidence of articles using Web-based research in American Psychological Association journals in 2003 and 2004 was relatively low (1.6%) (Skitka & Sargis, 2006). Skitka and Sargis classified studies that do use the Web for data collection into three types of Web-based psychological research:

1. *Translational research*, which involves adapting materials and methods originally developed for offline use for use on the Internet.
2. *Phenomenological research*, which focuses on the specific nature of how Internet use and Internet-based interaction (e.g., anonymous interaction) influence people’s thoughts, feelings, and behavior.
3. *Novel methodological use* of the Internet, such as use of information freely available on the net.

Among the translational research, Web surveys are having a profound effect on the survey research industry (Couper, 2000). The BBC, for example, commissioned a large-scale Web-based survey to investigate sex differences (Reimers, 2007). During 3 months of data collection, over a quarter of a million participants filled out the lengthy survey, which took about 40 minutes to complete.

An expanding theme for phenomenological research is the study of massive multiuser online role-playing games. Yee (2006), who used online survey data collected from 30,000 users over a 3-year period to explore gamers’ demographics, motivations, and derived experiences, suggested that these online environments could potentially be unique research platforms for the social sciences and clinical therapy, but that it was crucial to first establish that social behavior and norms in virtual environments are comparable to those in the physical world. In an observational study of the virtual community Second Life, Yee, Bailenson, Urbanek, Chang, and Merget (2007) found that social norms of gender, interpersonal distance, and eye gaze transfer into virtual environments even though the modality of movement is entirely different (i.e., via keyboard and mouse as opposed to eyes and limbs).

Massively multiplayer online games also constitute a platform for novel methodological use of the Internet, such as the study of cognition, learning and literacy in online digital contexts (Steinkuehler, in press). An online forum discussion of the online game *World of Warcraft*, for example, was used to evaluate the scientific mind habits used by the participants (Steinkuehler & Chmiel, 2006). In other research, Kafai, Feldon, Fields, Giang, and Quintero (2007) infected communities in Whyville.net, a teen online community with over 1.5 million registered players aged 8–16, with a virtual epidemic called Whypox, to understand the impact of the event on different aspects of community life and its potential as a model for educational interventions.
As access to and use of the Internet becomes more widely and representatively distributed globally, new opportunities exist for behavioral researchers to collect data online (Rhodes, Bowie, & Hergenrather, 2003). However, promising as it may be, Web-based research raises many methodological considerations. Six preconceptions that have been raised as likely limitations of Internet questionnaires were put to the test by Gosling, Vazire, Srivastava, and John (2004), who found that Internet samples are relatively diverse with respect to gender, socioeconomic status, geographical region, and age. They also discovered that Internet findings generalize across presentation formats, are not adversely affected by nonserious or repeat responders, and are consistent with findings from traditional methods (Gosling et al., 2004). Rhodes et al. (2003) concluded that many of the criticisms of online data collection are common to other survey research methodologies.

Here we use data gathered in a Web-based free-choice science-learning environment to learn more about these multidimensional interactions, using questions sent to an international Ask-A-Scientist site by people from various age groups and countries. Our study falls into the second and third categories of Skitka and Sargis’s (2006) classification of studies that use the Web for data collection: it is a phenomenological study in the sense that it compares girls’ interest in science in online and offline situations, and it is a novel methodological use of the Internet in that it makes use of information that is freely available on the net. Our research questions are

1. How did the percentage of questions asked by females change between countries, with age and with time?
2. How did the scientific interests, as they were reflected by the subject of the questions, change with age and time, between genders and between countries?

METHODOLOGY

Data Source

MadSci Network is an independent, award-winning nonprofit organization operating from a server in Scottsdale, AZ (http://www.madsci.org). It was founded in September 1995 as part of Washington University’s Young Scientist Program, a student-run organization in St. Louis dedicated to improving science literacy among K-12 students. Today, the MadSci Network receives 90–150 questions daily, most of which are answered automatically by the site’s search engine. Fewer than 20% of the questions are answered by nearly 800 globally distributed volunteering scientists, usually within 2 weeks.

Unlike most Ask-A-Scientist services (see further on), MadSci Network covers all branches of science, and does not focus on a specific subject area. It collects as much, and possibly more, information than most Ask-A-Scientist services, and stores key demographic information as metadata, making it easier to mine the information from the archives. Many other English-language Ask-A-Scientist services are available on the net, but none of them was found suitable for this study. The reasons for this were varied, among them: because they do not ask for the age of the asker (e.g., services run by Scientific American,1 the Internet Public Library,2 Yahoo! answers,3 and the paid service Google answer4), do not record all the information in their archives (e.g., Argonne National Labs5),

1 http://www.sciam.com/page.cfm?section=expertform
2 http://www.ipl.org/div/askus/
3 http://answers.yahoo.com/
4 http://www.answer.google.com/answers/
5 http://www.newton.dep.anl.gov/archive.htm
serve a limited age group (e.g., Ask Dr. Universe\textsuperscript{6} serves mostly elementary school children), have a rather small database (e.g., the service run by Cornell Center for Materials Research\textsuperscript{7} has collected just over 1,000 questions in the 8 years of its existence), or answer questions only on a specific topic (e.g., Howard Hughes Medical Institute\textsuperscript{8} receives only biology questions, Stanford University operates an Ask-a-Geneticist\textsuperscript{9} service that receives only genetics questions, and Ask Dr. Math\textsuperscript{10} obviously, deals with mathematics questions).\\

Sample Characteristics

Over 146,000 questions were sent to Madsci Network between 1996 and the first half of 2006. Almost 79,000 of the surfers disclosed their grade level, country of origin, and filled in the name and subject fields. These questions were used in our analysis. Even after this preliminary filtering, a few questions were missing some of the data, and therefore the \( n \) values differ between the variables.

The number of questions was not evenly distributed between years. In two cases, changes that were made in the site explain the decline in the number of questions relative to the previous year. For a full list of the number of questions for each year, see the appendix.

**Age Split.** 78,517 of the inquirers provided their grade level, and 66\% of the surfers were school students: 2.4\% were K-3 students, 10.5\% fourth–sixth graders, 26\% junior high school students, and 27.9\% senior high school students. Higher education undergraduates contributed 17.6\% of the questions, science graduates 7.5\%, and nonscience graduates 4.2\%. Teachers sent in 4\% of the questions.

**Gender Split.** Gender identification was based on the asker’s first name. Initial classification was done semiautomatically using an English name gender finder.\textsuperscript{11} Next, the names that were not automatically classified and appeared twice or more in the data (∼3,500 names) were analyzed individually using a baby name guesser,\textsuperscript{12} which operates by analyzing popular usage on the Internet. In this way, we were able to identify the gender of the asker in 48,360 of the questions. The rest were either names that could equally belong to boys or girls, meaningless scrambles, or names that appeared only once in the database. Of the gender-identifiable questions, 51.63\% were asked by boys (\( n = 24,968 \)) and 48.37\% were asked by girls (\( n = 23,392 \)).

**Split by Country of Origin.** 78,657 of the inquirers indicated their country of origin. The surfers originated from 143 countries that were grouped into 14 sociogeographic zones (Table 1), with 90\% of the questions originating from eight English-speaking countries (not necessarily as mother tongue): United States, UK, Canada, Australia, India, Singapore, Philippines, and New Zealand.

\textsuperscript{6} http://www.wsu.edu/DrUniverse/
\textsuperscript{7} http://www.ccmr.cornell.edu/education/ask/
\textsuperscript{8} http://www.hhmi.org/askascientist/
\textsuperscript{9} http://my.thetech.org/askScientist/askquestion.php
\textsuperscript{10} http://mathforum.org/dr.math/ask/submit.html
\textsuperscript{11} Japan Online Directory: http://epublishing.nademoya.biz/japan/names_in_english.php?nid=A
\textsuperscript{12} http://www.gpeters.com/names/baby-names.php
### TABLE 1
Distribution of Questions by Country of Origin

<table>
<thead>
<tr>
<th>Country Zone</th>
<th>Frequency</th>
<th>Percent</th>
<th>Contributing Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>52,594</td>
<td>66.7</td>
<td>United States, Puerto Rico</td>
</tr>
<tr>
<td>UK</td>
<td>5,437</td>
<td>6.9</td>
<td>UK (England, Scotland, Wales, Northern Ireland), Cayman Islands, British Virgin Islands</td>
</tr>
<tr>
<td>Canada</td>
<td>4,802</td>
<td>6.1</td>
<td>Canada</td>
</tr>
<tr>
<td>Australia and New Zealand</td>
<td>3,618</td>
<td>4.6</td>
<td>Australia, New Zealand</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>3,580</td>
<td>4.5</td>
<td>India, Pakistan, Nepal, Sri Lanka, Bangladesh, Burma</td>
</tr>
<tr>
<td>Far East</td>
<td>3,127</td>
<td>4</td>
<td>Singapore, Philippines, Malaysia, China, Hong Kong, Japan, Indonesia, Thailand, Taiwan, Korea, Vietnam, Papua New Guinea, Mongolia, Brunei</td>
</tr>
<tr>
<td>Western Europe</td>
<td>1,841</td>
<td>2.3</td>
<td>Ireland, Germany, Israel, Netherlands, France, Italy, Greece, Belgium, Switzerland, Spain, Portugal, Austria, Malta, Cyprus, Luxemburg, Gibraltar</td>
</tr>
<tr>
<td>Turkey, Iran, and the Arab World</td>
<td>1,024</td>
<td>1.3</td>
<td>Iran, Turkey, Egypt, United Arab Emirates, Lebanon, Oman, Saudi Arabia, Qatar, Palestinian authority, Kuwait, Iraq, Jordan, Bahrain, Morocco, Syria, Libya, Afghanistan, Algeria, Tunisia</td>
</tr>
<tr>
<td>Africa</td>
<td>832</td>
<td>1.1</td>
<td>South Africa, Nigeria, Ghana, Ethiopia, Mauritius, Kenya, Tanzania, Namibia, Uganda, Zimbabwe, Somalia, Mozambique, Malawi, Eritrea, Botswana, Gambia, Rwanda, Lesotho, Zambia, Chad</td>
</tr>
<tr>
<td>Latin America</td>
<td>744</td>
<td>0.9</td>
<td>Brazil, Mexico, Barbados, Nicaragua, Colombia, Argentina, Peru, Venezuela, Chile, Guyana, Costa Rica, Guatemala, Honduras, Dominican Republic, Ecuador, Panama, Bolivia, Paraguay, El Salvador, Belize, Saint Lucia, Antigua &amp; Barbuda, Dominica, Uruguay, Venezuela</td>
</tr>
<tr>
<td>Northern Europe</td>
<td>628</td>
<td>0.8</td>
<td>Sweden, Norway, Denmark, Finland, Iceland</td>
</tr>
<tr>
<td>Former Communist Countries</td>
<td>296</td>
<td>0.4</td>
<td>Russia, Romania, Poland, Croatia, Hungary, Lithuania, Bulgaria, Latvia, Slovenia, Estonia, Albania, Czech Republic, Kazakhstan, Macedonia, Serbia &amp; Montenegro, Ukraine, Slovakia, Belarus, Yugoslavia</td>
</tr>
<tr>
<td>English-Speaking Caribbean</td>
<td>107</td>
<td>0.14</td>
<td>Jamaica, Trinidad &amp; Tobago, Bahamas, Saint Kitts &amp; Nevis, Bermuda</td>
</tr>
<tr>
<td>Pacific Ocean Islands</td>
<td>37</td>
<td>0.05</td>
<td>Cook Islands, Marshall Islands, Samoa, Guam</td>
</tr>
</tbody>
</table>

*a209 entries were classified as “unknown.”

*bIn bold: more than 500 entries, in italic: 100–500 entries.

*cNo longer exists.
The number of questions arriving from the different countries was found to be weakly, but significantly, correlated to the country’s gross domestic product (GDP) per capita (n = 120 countries,\textsuperscript{13} \( p = .006 \)) and to its number of Internet users per 1,000 people (n = 85 countries,\textsuperscript{13} \( p = .04 \)) according to data from the 2005 \textit{Human Development Report} (United Nations Development Programme, 2005).

**Subject of Questions.** The questions were allocated by the surfers into one of 25 topics. For clarity, we pooled them into seven main fields of interest, appearing here in their order of popularity: biology (e.g., How long does it take for the calories in your body to transform into fat?), chemistry (e.g., Will a coke fizz if it is opened in equal pressure?), physics (e.g., If I want to find out my volume using a bath tub, how do I do it?), technology (e.g., Can you give me any links to research on language translation by inserting a chip into a person’s brain?), earth sciences (e.g., Why can’t Hawaii get rid of its trash by putting it in a volcano?), astronomy (e.g., Why do meteorites in the same orbit enter the atmosphere at different angles?), and science history (e.g., When did scientists realize that the brain, not the heart, was used for thought?).

Biology includes the following Madsci topics (in order of popularity): biochemistry, general biology, zoology, botany, anatomy, cell biology, environment and ecology, medicine, genetics, microbiology, neuroscience, agricultural sciences, evolution, molecular biology, development, virology, immunology, and biophysics. Technology includes the Madsci topics engineering and computer science. The other topics were not further subdivided by the Web site’s operators.

Although all of the questions in this sample were self-generated by the askers, it is important to note that some of them were school related. These questions were not spontaneously raised by the students, but were the consequence of a school assignment. In the current study, we did not identify these school-related questions. However, in a previous study that rigorously examined a 3-month sample of 4th–12th graders’ questions sent to the same Web site, questions were classified as “school related” if the question explicitly stated that the information is required for a school assignment, such as a science-fair project, report, or homework (e.g., Who wrote the \textit{Origin of Species}? This is for a bonus question) (Baram-Tsabari et al., 2006). In that analysis, we learned that chemistry is characterized by a relatively large number of school-related questions, while this is less true for astrophysics. Biological topics such as anatomy and physiology, sickness and medicine, and genetics and reproduction were all characterized by relatively more “spontaneous” than school-related questions. Botany and mycology, microbiology, virology, and cell biology yielded many more teacher- and textbook-generated questions than spontaneous ones. Topics such as ecology and neurology were almost equally distributed among the two question types (Baram-Tsabari et al., 2006).

**Statistical Analysis.** Unless otherwise indicated, a two-tailed Pearson chi-square test was used to calculate probabilities. Not all of the inquirers provided their full details; therefore, sample sizes differ from graph to graph and are indicated by \( n \) values. Significant differences within proportions were determined according to a cell chi-square test.

**FINDINGS**

A decade of questions sent to an Ask-A-Scientist Internet site, which were analyzed by age, gender, country of origin and the time they were sent, provided answers to questions

\textsuperscript{13} Number of countries that appeared in the sample and had the relevant data in the report.
concerning the level of female participation and topics of interest in science as they were reflected by the subjects of the questions.

**Age and Geographical Distribution of Female Participants**

Overall, among the gender-identifiable questions, 51.63% were asked by boys \( n = 24,968 \), 48.37% by girls \( n = 23,392 \). However, among K-12 students, we found a dominance of female contributors (see further on). Different countries displayed different female participation patterns \( p < .0001 \).

Female participation rates in 39 countries that contributed 25 or more gender-identified questions were analyzed \( n = 47,749 \). The most “feminine country,” with a female participation of 61.1%, was Egypt, followed by Hong Kong, Colombia, Philippines, Switzerland, Iran, and Indonesia, which all presented more female contributions than male ones. The United States, Ireland, Australia, and New Zealand had an almost equal gender split, while the remaining countries all presented male dominance. Completing the list was Sweden with only 5.6% of the contributions sent by girls, and Peru, the Netherlands, Brazil, Romania, and Denmark were close behind (Figure 1).

Female participation was found to be correlated \( n = 40 \) countries, \( r = -.36, p = .02 \) with the difference between males’ and females’ science scores in the eighth grade among countries who had participated in the “Trends in International Mathematics and Science Study” (TIMSS) research (Martin et al., 2004). A correlation was not found, however, with the gender-related development index of the UN14 (United Nations Development Programme, 2005), indicating that the female participation in this free-choice science activity was not related to the level of equity in the different countries.

A significant difference \( p < .0001 \) was found in the female participation among the different age groups \( n = 46,578 \). Overall, females sent the majority of contributions among K-12 students, but their percentage dropped upon moving from junior high to high school—they became the minority among undergraduates and even more so among graduates (Table 2, last row).

One might conclude that the decrease in females’ contributions with age was due to their loss of interest and not to an increase in male interest, since the decrease in the percentage of girls’ contributions was accompanied by a decrease in their absolute number, while the increase in percentage of boys’ contributions was also accompanied by a decrease in their absolute number. This decrease in absolute numbers was due to a general decline in the number of questions sent by the older age groups: 26% were sent by junior high school students, 27.9% by senior high school students, 17.6% by undergraduates, and only 11.7% by graduates. Note that we are not suggesting that students tend to lose interest in science as they grow older on the basis of the drop in the absolute number of questions in our sample, since a specific Internet site, such as MadSci, has a target audience, and is not equally appealing to every age group. Moreover, the older people may have found other sources for answering their questions.

The decrease in female participation with age was observed all over the world. This pattern was clearly displayed for questions arriving from the United States, UK, Canada, Australia, and New Zealand (Table 2). Girls from the Far East dropped their participation.

---

14 Human development reports, Technical Note 1: The gender-related development index (GDI) adjusts the average achievement of a country to reflect the *inequalities* between men and women using three basic dimensions of human development: (1) A long and healthy life, as measured by life expectancy at birth. (2) Knowledge, as measured by the adult literacy rate (with two-thirds weight) and the combined primary, secondary, and tertiary gross enrollment ratio (with one-third weight). (3) A decent standard of living, as measured by GDP per capita.
Figure 1. The percentage of female contributors of questions from various countries ($n = 47,749$). The $x$ axis passes at 50%. Only countries that had 25 gender-identifiable contributions or more are shown. In brackets: the number of contributions from the specific country, number of gender-identified contributions. Difference between countries is significant at $p < .0001$. 
<table>
<thead>
<tr>
<th>Sociogeographic Zone</th>
<th>K-3 (n = 1,089)</th>
<th>4–6 (n = 5,132)</th>
<th>Junior High School (n = 13,140)</th>
<th>Senior High School (n = 13,401)</th>
<th>Undergraduates (n = 8,475)</th>
<th>Graduates (n = 5,236)</th>
<th>Significance</th>
<th>Average Female Participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States (n = 33,604)</td>
<td>56.98</td>
<td>59.54</td>
<td>58.46</td>
<td>51.55</td>
<td>40.44</td>
<td>23.85</td>
<td>***</td>
<td>51.00%</td>
</tr>
<tr>
<td>Australia and New Zealand (n = 2,168)</td>
<td>47.22</td>
<td>67.07</td>
<td>60.25</td>
<td>56.07</td>
<td>35.93</td>
<td>27.07</td>
<td>***</td>
<td>50.60%</td>
</tr>
<tr>
<td>Canada (n = 2,825)</td>
<td>51.28</td>
<td>51.64</td>
<td>64</td>
<td>50</td>
<td>38.25</td>
<td>27.24</td>
<td>***</td>
<td>49.59%</td>
</tr>
<tr>
<td>Turkey, Iran, and the Arab World (n = 259)</td>
<td>na</td>
<td>na</td>
<td>73.68</td>
<td>51.09</td>
<td>55</td>
<td>43.56</td>
<td>*</td>
<td>49.42%</td>
</tr>
<tr>
<td>UK (n = 3,569)</td>
<td>51.19</td>
<td>50.49</td>
<td>57.87</td>
<td>51.63</td>
<td>42.64</td>
<td>21.16</td>
<td>***</td>
<td>44.21%</td>
</tr>
<tr>
<td>Far East (n = 1,234)</td>
<td>69.57</td>
<td>49.06</td>
<td>40.08</td>
<td>42.73</td>
<td>37.41</td>
<td>35.81</td>
<td>*</td>
<td>40.84%</td>
</tr>
<tr>
<td>Western Europe (n = 877)</td>
<td>16.67</td>
<td>42.11</td>
<td>47.5</td>
<td>49.54</td>
<td>33.5</td>
<td>32.99</td>
<td>**</td>
<td>37.86%</td>
</tr>
<tr>
<td>Latin America (n = 383)</td>
<td>na</td>
<td>61.54</td>
<td>61.76</td>
<td>40.46</td>
<td>20</td>
<td>19.84</td>
<td>***</td>
<td>32.11%</td>
</tr>
<tr>
<td>Southeast Asia (n = 695)</td>
<td>na</td>
<td>na</td>
<td>48</td>
<td>47.33</td>
<td>27.31</td>
<td>27.03</td>
<td>**</td>
<td>31.94%</td>
</tr>
<tr>
<td>Africa (n = 297)</td>
<td>na</td>
<td>na</td>
<td>58.62</td>
<td>42.11</td>
<td>24.07</td>
<td>13.75</td>
<td>***</td>
<td>29.29%</td>
</tr>
<tr>
<td>Northern Europe (n = 381)</td>
<td>na</td>
<td>na</td>
<td>56.25</td>
<td>3.21</td>
<td>14.75</td>
<td>10.24</td>
<td>***</td>
<td>10.24%</td>
</tr>
<tr>
<td>All Sample (n = 46,473)</td>
<td>54.91</td>
<td>58.83</td>
<td>58.45</td>
<td>51.24</td>
<td>38.21</td>
<td>25.23</td>
<td>***</td>
<td>48.89%</td>
</tr>
</tbody>
</table>

*Only sociogeographic zones that had more than 200 gender-identified contributions were included in this analysis.

**Including the less represented zones that do not appear in the table.

***Significance of gender vs. age group comparison. Marked: ***, p < .0001, **, p < .001, *, p < .05.

†Percentage does not include teachers and contributors who did not disclose their age group.

‡Age groups that had fewer than 10 gender-identified contributions were not included in this analysis. Na = not analyzed.
rates sooner: from almost 70% among the K-3rd graders to 49% among fourth–sixth graders, 40%–43% in high school, and 37%–35% among undergraduates and graduates. The same happened in Latin America, where the female majority among fourth–ninth graders became a minority in high school (40%), and even more so among undergraduates and graduates (20%).

In Southeast Asia, the female participation rate for 7th–12th graders was 47%–48% and 27% for undergraduates and graduates. In Africa, girls’ participation gradually fell from 59% in junior high school to 14% among graduates. The same was found with data from Turkey, Iran, and the Arab world, where girls’ participation gradually fell from 74% in junior high school to 44% among graduates. In Western Europe, girls’ participation grew as long as the girls were K-12 students, reaching almost 50%, and then fell to 32%–33% among undergraduates and graduates. In Northern Europe, girls contributed 56% of the high school questions, but only 3% of the undergraduate and 15% of the graduate questions. The remaining sociogeographical zones did not provide enough data to allow this kind of analysis.

The rate of female participation changed between the years 1996 and 2006 ($p < .0001$); however, a constant trend was not found, and females’ percentages fluctuated between years (1996, 36.9%; 1997, 44.7%; 1998, 52.4%; 1999, 42.6%; 2000, 47.6%; 2001, 47.1%; 2002, 50.8%; 2003, 51.9%; 2004, 49.8%; 2005, 49.2%; 2006, 45.4%). A constant trend was not identified in any of the separate sociogeographical zones either.

**Identifying Scientific Interests**

Overall, the questions referred to the following scientific disciplines, appearing here in their order of popularity: biology (42.5%), chemistry (19.1%), physics (17.9%), technology (7.2%), earth sciences (6.4%), astronomy (5.6%), and history of science (1.4%).

*Science Education*
As expected, a significant difference \( (p < .0001) \) was found in the distribution of question topics between boys and girls \((n = 42,705)\). Boys were more interested than girls in physics (23.5% vs. 11% of the boys’ and girls’ questions, respectively) and technology (9.2% vs. 3.2% of the boys’ and girls’ questions, respectively), while girls were more interested than boys in biology (51.1% vs. 36.6% of the girls’ and boys’ questions, respectively) (Figure 2). Statistical differences between girls and boys were found in all subjects, but those which contributed the most to the differences were physics, technology, and biology.\(^{15}\)

Question topic differed significantly \( (p < .0001) \) between age groups \((n = 69,529)\). Interest in physics and technology increased with age (physics: from 13% among K-3 to 24% among graduates; technology: from 4% among K-3 to 15% among graduates), while interest in the earth sciences decreased with age (from 9.9%–13.3% among K-6 to 3.7% among graduates). Interest in chemistry peaked at 25% during the high school years, probably due to the relative abundance of school-related questions in this subject among this age group (Baram-Tsabari et al., 2006).

The two genders developed different scientific interests as they grew: distinctively different trends were identified in the development of interest in biology and physics \((n = 41,028)\). Males developed interest in physics with age (from 16.3% of the questions among K-3 to 30.5% among graduates), while females did not seem to develop such an interest to the same degree (9.8% of the questions among K-3 and 12.4% among graduates). Males lost some interest in biology with age (from 45.5% of the questions among K-3 to 31.1% among graduates), while females actually showed more interest in the subject with age (from 49.5% of the questions among K-3 to 60.3% among graduates) (Figure 3).

The initial gap in the interests of K-3 boys and girls applies only to physics, but as the children mature, the gap between their stereotypically gendered science interests widens and applies to biology as well.

A significant difference \( (p < .0001) \) was found in the distribution of the questions topics with years \((n = 69,869)\). Some of this difference might originate from a decrease in interest in astronomy (from 10.8% in 1996 to 4.8% in 2006) and an increase in interest in technology (from 4.1% in 1996 to 8.4% in 2006). The rest may originate from nondirected fluctuations between the years that were found to be significant only due to the massive sample size. Changes in question topics across years could be connected to science-related current events, such as hurricanes. In such cases, the site operators provide a link to frequently asked questions on the subject in a visible place on the homepage.\(^{16}\) However, when we specifically looked for an increase in questions about tsunamis after the deadly tsunami hit countries in Southern Asia in December 2004, we found only a few questions on the topic.

We found a significant difference \( (p < .0001) \) in the distribution of question topics between different geographical zones, with the biggest differences between the United States and countries in Southeast Asia. Since the female participation at this site was very different in these two geographical zones (51% in the United States vs. 32% in Southeast Asia), the distribution of question topics was compared for each gender separately. Regardless of the large stereotypical gender-related differences in interests, both male and female Americans were more interested in chemistry, earth sciences, and astrophysics than male and female students from Southeast Asia, but less interested in physics and technology than their overseas peers.

\(^{15}\) According to a cell chi-square test.

\(^{16}\) http://www.madsci.org/FAQs/earth/hurricanes.html

Science Education
Almost 79,000 questions sent over the course of a decade to an international Ask-A-Scientist site were used to learn about the scientific interests of boys and girls of different age groups from various countries in an online free-choice science-learning environment. The site was found to serve as a Web-based bypass for traditional gender inequities in science education, while maintaining the usual gender gap in interest in biology and physics.

One finding that emerged from the analysis was the absence of correlation between the gender-related developmental index of the UN and the level of female participation among the contributions originating from different countries. In other words, countries that emphasize equity in their education policy and legislation did not have a higher percentage of girls sending questions to this science site than countries which do not promote gender equity.

*Science Education*
Gender equity has been a dominating political and educational concern in Scandinavian countries since the mid-1970s (Sjøberg, 2000). For example, for the past 20 years, Sweden has launched general information campaigns aimed at broadening girls’ educational and occupational choices and stimulating their interest in science and technology, promoted intervention projects for teachers and science centers, reviewed girls’ attitudes and issued relevant policies. In Denmark, among other initiatives, the government has placed equal-opportunity consultants in employment offices, and groups of women science teachers have been linked to tertiary institutions (Harding & Parker, 1995). At the same time, in many Muslim countries, gender-based discrimination, coupled with social and cultural barriers, limits women’s access to and participation in higher education, and career opportunities for female science graduates are more limited than for their male counterparts (Hassan, 2000).

Nevertheless, Egypt, Iran, and Indonesia (ranked 119, 99, and 110, respectively, in the UN’s gender-related development index [GDI]) displayed a female dominance among contributors, while Sweden, Denmark, and the Netherlands (ranked 6, 14, and 12, respectively, in the GDI) had among the lowest percentage of female contributors in our research. The relationship between the GDI and female participation is neither linear nor inverted—it simply does not exist. For example, the Philippines (ranked 84 in the GDI) displayed a female majority, while Peru (ranked 79 in the GDI) displayed only a small minority of girls.

It might prove fruitful to view these findings in light of the results from the international project “Science and Scientists” (Sjøberg, 2000), which found that children in rich and technologically developed countries show a low or moderate interest in learning science topics, with Japan, Norway, and Sweden being the lowest. Children from developing countries, on the other hand, appear to be very interested in learning science. The gender profile is also intriguing: in most developed countries, boys are more interested in learning science than girls, while in most developing countries, the opposite is true (Sjøberg, 2000). Sjøberg tentatively explained this pattern with the idea that obtaining an education is a luxury in developing countries, especially for girls, while it is perceived as a burden by many students in developed countries. This finding was repeated in results from the ongoing international “Relevance of Science Education” [ROSE] project which found that the higher the level of development in a country, the lower the level of interest expressed by students toward learning about science- and technology-related topics and having a career in those fields (Sjøberg & Schreiner, 2005). Sjøberg and Schreiner suggest that this pattern might reflect the postmaterialistic values of youth in developed societies, in contrast to the view of science in developing countries as a key for improving the quality of life. Schreiner (2006) interpreted interest as a sign of late-modern identity, and explained that the more modernized a country, the more its girls accentuate their femaleness and boys their maleness.

One could suggest that if a school in a developed Western country provides a richer environment for science learning, may be girls just do not need the Ask-A-Scientist tool. However, this does not explain why boys from the same country do need the tool. Moreover, research tells us that gender differences are not smaller in technologically advanced countries or in advantaged socioeconomic groups with regard to percentage of female researchers (UNESCO Institute for Statistics, 2006) and that gender differences in science achievements are higher among school students (Martin et al., 2004). We can further hypothesize that females in developed countries have a wider range of educational and occupational possibilities, and therefore do not view science as a unique escape route from their traditional gender roles. It is also possible that females’ interest in science is a product of their wish to impact society, or even a form of rebellion against a limiting society, similar to the way in which forming a reading group of English literature served the participating women in “Reading Lolita in Tehran” (Nafisi, 2003).
The case of Iran presents the most fascinating paradox: on the one hand, postrevolutionary educational policy in Iran is characterized by the banning of coeducation, compulsory veiling of female students over the age of 6, explicit gender stereotyping in school textbooks, guiding female students toward feminine specializations and occupations (such as sewing, nursing and teaching), and creating a traditional atmosphere in schools to educate “modest girls and courageous boys” (Mehran, 2003). The republic’s education plan assigns different roles and responsibilities to boys and girls, embedded in principles such as: “The Iranian educational system should recognize the identity of a woman and her role in the family and the society on the basis of Islam and plan for the content and method of her schooling accordingly,” and “… their [girls’] vocational guidance should take into consideration the kinds of occupations needed by women, best fulfilled by women, or most fitting their role and responsibility in the family.” On the other hand, since the Islamic revolution, there has been a significant increase in female enrollment at every educational level, including university (Mehran, 2003), Iran presented no gender gap in science achievements in the last TIMSS research (Martin et al., 2004), and it is among the few countries in our study that displayed a female majority among contributors of science questions.

This paradox may indicate that the Internet as a free-choice science-learning environment has a potentially empowering and democratic role that is especially relevant to populations which are deprived of equal opportunities to learn formal science. Indeed, according to estimations, Internet use in Iran is more common than in any other country in the Middle East except for Israel, and Farsi is the fourth most prevalent language on the Internet (Menshari, 2007). Rahimi (2003) argues that the Internet, as a new and advancing means of communication, has played an important role in the ongoing struggle for democracy in Iran, allowing Internet users—especially women—to take advantage of the freedom provided by the Internet as an alternative medium for the expression that is denied them in real public spaces.

Inequitable social distribution of knowledge and access to knowledge is not merely a phenomenon of non-Western cultures—such inequities exist within Western cultures as well (Kyle, 1999). The Internet seems to provide an attractive science-learning environment for female students, who are traditionally found to be less interested in science than males in offline situations, but constitute the majority of contributors among the K-12 age groups in our sample. A female majority has been found in other Web-based free-choice science-learning environments, such as Whyville (Aschbacher, 2003), KidsConnect (Lankes, 2003), Ask-A-Geneticist (our unpublished results), and Argonne National Labs (our unpublished results), but not in Ask-A-Scientist TV shows for kids (Baram-Tsabari & Yarden, 2005) or adults (Baram-Tsabari & Yarden, 2008). The difference in female participation in Web-based and TV-based settings might be a product of the level of anonymity and safe atmosphere they provide. Furthermore, the textual modality of the Web-based Ask-A-Scientist answer is better suited to lengthy explanations and essay-like answers that demand a deep understanding, rather than short factual “right answers” that are known to be less female friendly (Behling, 2002). Previous studies on students’ questions have found that on average, girls are more interested in asking for explanations than boys, while boys are more interested in factual and methodological information than girls (Baram-Tsabari & Yarden, 2005, 2008). It is interesting to note that we did not find an increase in female participation over the decade in which the questions were asked, which may mean that the Internet as a science-learning setting has not become more female friendly over the years.

As attractive and inclusive as it may be, girls seem to lose their wish to submit science questions to Web sites as they grow older. This finding mirrors previous research in which American girls’ attitude to science was found to become increasingly negative with age.
It is possible that the core of the decrease in female contributions with age is gendered philosophies or perspectives, rather than science interest specifically. Schreiner and Sjøberg (2007), for example, suggest that young Westerners, especially girls, do not want to have an identity that is seen to be connected with being a physicist or an engineer. Such a tendency might constitute a reason for losing interest in science. In this study, we were unable to identify gender-related and gender-unrelated reasons for losing interest in science. It is important to note, however, that the decrease in female contributions in our sample occurred at the same time as that which is expected from school-based findings, where girls were found to lose interest in science upon moving from junior to senior high school and into higher education.

We found the polar pattern of girls’ higher interest in biology and boys’ greater interest in physics in free-choice science-learning settings to be similar to the situation described within formal science education. However, this seemingly spontaneous interest might be to some extent a result of formal schooling, and not a completely independent measure of out-of-school interest.

However, an interesting disagreement was found regarding the development of interest in physics. A study conducted in German schools reported that girls find physics as a school subject less and less interesting as they grow older, while boys do not lose their interest in it (Hoffmann, 2002; Hoffmann & Haussler, 1998). Based on our data, it seems that girls do not lose their interest in physics, but simply never develop such an interest in the first place.

Using our current analysis, it is impossible to separate the apparent connection between course-taking patterns, historically noted in the literature, and interests as they are reflected by self-generated questions. However, this finding is aligned with results obtained in a former study that used only spontaneously asked (non-school-related) questions from three different databases that provided Ask-A-Scientist applications for different age groups in different countries and languages (Baram-Tsabari & Yarden, 2008). The seemingly contradictory findings obtained in German schools may result from the different settings and methodologies used for gathering the data. We suggest that this inconsistency be further studied.

**METHODOLOGICAL CONSIDERATIONS AND LIMITATIONS**

Pursuant to a discussion of our findings, it is important to acknowledge some limitations of the research methodology and of the uncontrolled sample that was used in this study.

Coverage error is presently the biggest threat to inferences from Web surveys, at least to groups beyond those defined by access to or use of the Web (Couper, 2000). Coverage error is a function of the mismatch between the target population (the set of people one wishes to study) and the frame population (the proportion of the target population that can potentially be reached via the Web) (Couper, 2000). However, the most recent Pew Internet Project survey found that 87% of all American youth between the ages of 12 and 17 use the Internet (Rainie & Hitlin, 2005), and in the fall of 2003, nearly 100% of public schools in the United States already had access to the Internet (NCES, 2005), theoretically allowing all students to send their questions and be part of our sample.

However, use of the Web has linguistic and socioeconomic aspects as well. To state one example, a sample of the geographical distribution of visitors to a science-controversies online forum that deals with genetically modified food showed heavy participation from the United States. The study’s authors assumed that this was a reflection of the number of people with Internet access and fluency in English rather than an interest in debates about
food (Triunfol & Hines, 2004). We assume that the heavy participation of users originating from English-speaking countries, particularly the United States, is in no way representative of the level of interest in science in different countries.

The nature of our data poses some methodological problems as well. In our dataset, the askers are the ones who allocated the question to a certain subject area, and we cannot guarantee that this topic classification would be identical to a classification performed by a science-education researcher. We attempted to maximize the reliability of the data by including in the analysis only questions from participants who provided background data regarding their name, age, and origin. Our intention was to exclude “casual” participants who just happen to click with no formulated question in mind, and saboteurs, who just want to send meaningless scrambles or annoy the site operators (and the scientific community in general). However, there may be a sampling bias in using only data from those who report their real names as opposed to meaningless scrambles.

The self-reporting nature of the dataset is actually one of the strengths of this study. A problem of surveys in general is measurement error—the deviation of the responders’ answers from their true values on the measure (Couper, 2000). Traditional questionnaire-based interest studies rely on self-reporting to measure interest. This study, however, does not ask students to respond to our requests, but observes cases in which they initiate the action themselves, for their own purposes.

The self-selecting sample used in this research does not represent all children in a particular country. It represents a group of children who might be more interested in science and have easier access to resources than the child population as a whole. There is a marked difference in the access of children from different socioeconomic statuses to the Internet, which was our source for the questions. Furthermore, students who are not motivated to learn science are probably not represented in this self-selecting sample at all, regardless of their socioeconomic status. Other children that may be very interested in science but do not send questions are also not represented. Therefore, the opportunistic nature of the sample places some constrictions on the validity of our results.

The validity of the study can be supported by the notion of using data that originate from the researched population itself, not as a response to a stimulus from a researcher, thus ensuring high ecological validity. However, it is important to note that although students generated all of the questions in our sample, not all of the questions were the outcome of an intrinsic motivation to know. Some of the questions were required for school assignments and were originally raised by teachers or textbooks. This setback is even more salient in studies that examine students’ questions in a classroom setting. For a list of school-related versus spontaneous topics, see Baram-Tsabari et al. (2006).

Another way to achieve validation is by comparing any conclusions drawn with other independent observations: using an informal data source and a new methodology, our results confirm and reinforce revelations made using the traditional questionnaire methodology in a formal setting. This agreement with findings described in the literature, which were gathered using control samples, also serves to bolster confidence in our new, never before described findings. Therefore, we assume that the trends described here represent, to a certain degree, the interests of many children, and that posing questions represents a measure of student orientation toward science interest. It is also important to note that in formal settings, some students are also more likely to ask questions than others, with lower achieving students becoming less active participants with age (Good et al., 1987). Reliability may be assured by the use of a very large sample (Reid, 2006).

Although Web-based experiments of the kind used here are more difficult to control in some respects than those conducted in a classroom setting, they present an important methodological advantage for studying students’ self-guided science learning, taking into
consideration that this kind and amount of data does not exist anywhere outside the Web. Therefore, this methodology is better suited to studying the dynamic educational reality of the last decade.

This research was supported by the Sacsta Rashi Fellowship for honor students in science education. The authors would like to thank Mrs. Yetty Varon and Dr. Hillary Voet for their expert statistical advice, Mrs. Rakefet Halevi for her valuable help in the coding of the questions, and Professor Edgar Jenkins and the three anonymous reviewers for their insightful remarks on an earlier draft of the manuscript. AY is the incumbent of the Helena Rubinstein Career Development Chair.

**APPENDIX**

**Distribution of Questions Sent to Madsci Network Since Its Establishment**

<table>
<thead>
<tr>
<th>Year</th>
<th>Questions Submitted</th>
<th>Questions Used for the Analysis</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>21</td>
<td>–</td>
<td>The pilot year. Data were not included in the analysis.</td>
</tr>
<tr>
<td>1996</td>
<td>1,165</td>
<td>961</td>
<td>More than 5,000 of the surfers did not disclose their age.</td>
</tr>
<tr>
<td>1997</td>
<td>8,002</td>
<td>1,659</td>
<td>More than 10,000 of the surfers did not disclose their age.</td>
</tr>
<tr>
<td>1998</td>
<td>14,327</td>
<td>3,047</td>
<td>More than 10,000 of the surfers did not disclose their age.</td>
</tr>
<tr>
<td>1999</td>
<td>20,088</td>
<td>6,133</td>
<td>More than 10,000 of the surfers did not disclose their country of origin.</td>
</tr>
<tr>
<td>2000</td>
<td>25,165</td>
<td>10,438</td>
<td>More than 10,000 of the surfers did not disclose their country of origin.</td>
</tr>
<tr>
<td>2001</td>
<td>18,372</td>
<td>13,291</td>
<td>In the fall of 2001, an additional page was placed before the Ask-A-Scientist form and all questions were run through the local search engine before writing to the site.</td>
</tr>
<tr>
<td>2002</td>
<td>9,659</td>
<td>7,016</td>
<td>At the beginning of the year, a search engine form was placed on every page, and question volume dropped by ~50%. The questions were collected until June 15.</td>
</tr>
<tr>
<td>2003</td>
<td>14,947</td>
<td>11,205</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>14,802</td>
<td>10,888</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>15,667</td>
<td>11,171</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>4,314</td>
<td>3,056</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>146,529</td>
<td>78,865</td>
<td></td>
</tr>
</tbody>
</table>

**REFERENCES**


Science Education


Science Education


QUESTIONS ANALYZED BY AGE, GENDER, AND COUNTRY


Science Education


