Authors' final version of: Cohn, S. and Fraser, B. 2015. Effectiveness of student response systems in terms of learning environment, attitudes and achievement. Learning Environments Research. http://doi.org/10.1007/s10984-015-9195-0

EFFECTIVENESS OF STUDENT RESPONSE SYSTEMS IN TERMS OF LEARNING ENVIRONMENT, ATTITUDES, AND ACHIEVEMENT

Abstract In order to investigate the effectiveness of using Student Response Systems (SRS) among grade 7 and 8 science students in New York, the How Do You Feel About This Class? (HDYFATC) questionnaire was administered to 1,097 students (532 students did use SRS and 565 students who did not use SRS). Data analyses attested to the sound factorial validity and internal consistency reliability of the HDYFATC, as well as its ability to differentiate between the perceptions of students in different classrooms. Very large differences between users and non-users of SRS, ranging from 1.17 to 2.45 standard deviations for various learning environment scales, attitudes and achievement, supported the efficacy of using SRS.

Keywords Achievement, Attitudes, Learning environment, Middle-school science, Student Response Systems (SRS)

Introduction

A recent trend involves using Student Response Systems (SRS) to track student responses during class activities. Because new technology, such as SRS, requires funds to procure, implement and maintain, it is important to evaluate its effectiveness in terms of fostering student outcomes as well as creating a positive learning environment. Although past research on SRS has involved the university level and using informal questionnaires and interviews, few formal in-depth evaluation studies have been conducted at the middle-school level. Therefore we undertook this study to fill gaps in research knowledge about the effectiveness of using SRS in terms of the learning environment of science classes, students' attitudes towards these classes and how comfortable they feel there. Our research questions were:

1. Is the How Do You Feel About This Class? questionnaire valid and reliable with 7th and 8th grade science students?
2. Is the use of Student Response Systems effective in terms of:
   a) the learning environment?
   b) student attitudes?
   c) student achievement?

Student Response Systems

Student Response Systems (SRS), also commonly referred to as Personal Response Systems (PRS) and other names, provide communication between individual students and the teacher. Each student has a transmitter, often called a 'clicker', which resembles a small television remote control (Draper and Brown 2004). When students respond to a question, they aim their transmitter at a sensor, which is connected to a computer. Running on this computer is a program that tabulates the results, which are projected onto a screen for all students to see.

Questions can be presented to the class orally, projected onto a screen, or provided on sheets of paper. When students respond, their answers are anonymous. However, the instructor has the option of copying each student's identification number from his/her transmitter so that he/she can see each student's results. When the responses are projected for the class to view, none of the students know how each other responded.

Technology similar to SRS was used in the 1960s and 1970s based on hardwired systems that were typically made in-house. These systems were more cumbersome than the modern technology. Even though students reported positive attitudes towards these devices, research provided no evidence of a measureable advantage over regular classroom instruction (Judson and Sawada 2002). Casanova (1971) even suggested that students had positive attitudes that led to decreased attrition, which lowered the overall mean scores compared to the control group. These early devices were typically used for quizzes to provide answers to questions to the instructor, with very little class discussion about the results. Littauer (1972) reported that these devices ‘accidentally’ fostered student collaboration.

Most research on the effectiveness of SRS has been at the college level. It has been noted that, in many classrooms in which Student Response Systems are used, more class discussion takes place. Often, teachers allow students to discuss the answers to questions when there is a diversity of answers obtained through use of the clickers. Peer instruction raises interest and enjoyment in science (Duncan 2005). These discussions usually move the group towards the correct answer and student understanding increases. Not only do the students understand the information better, but their retention increases as well (Duncan 2006; Mazur 1997). When a political science professor used SRS, there was an improvement in the quality of discussions (Guess 2008). In a study undertaken at the University of Wisconsin, most students responded with ‘agree’ and ‘strongly agree’ when asked if clickers led them to be more engaged in class and increased the frequency of their participation in the course (Joosten and Kaleta 2006).

In courses in which clickers are used to record attendance, a marked increase in attendance has been reported (Draper and Brown 2004). In a study undertaken in a Statistics for Psychologists course at the University of Glasgow, attendance prior to the use of the clicker system was around 32% and, after implementing SRS, it increased to around 57% (Wit 2003). Guess (2008), Burnstein and Lederman (2001), Duncan (2008) and Homme, Asay and Morgenstern (2004) have all reported an increase in class attendance in their research.

Improvements have also been noted in examination scores. At Ohio State University, final examination scores were about 10% higher for classes in which SRS were used when compared with classes in which this technology was not used (Guess 2008). Mazur (1997) found that there were dramatic increases in pre–post gains in students' knowledge in physics classes when SRS were used compared with students who did not use SRS. In one of the most comprehensive studies on SRS, the pass rate for students who used SRS was approximately 50% greater than for students who did not use these systems. Also, the standard deviation for the SRS groups was substantially lower than for the non-SRS groups, suggesting a more consistent understanding among students who used SRS (Poulis, Massen, Robens and Dilbert 1998). Martyn (2007) also reported that scores were consistently higher for students who used clickers.

Student attitudes towards their classes in which SRS were used also improved (Duncan 2008; Martyn 2007). In a general science chemistry class, 90% of the students responded with 'useful' or 'very useful' when asked how they rated the use of clickers in the class (Guess 2008). Most students also responded with 'agree' or 'strongly agree' when asked to respond to a statement about whether the use of SRS was associated with class interest and enjoyment (Duncan 2008). Most students responded with 'agree' or 'strongly agree' when answering about whether they were happy with using clickers. The same was seen when students were asked if they would take another course that involved SRS (Joosten and Kaleta 2006). Students reported that using SRS was 'fun' (Roberts 2005; Siau, Sheng and Nah 2006). Numerous other studies support these improvements in student attitudes when SRS were used (Burnstein and Lederman 2001; Crouch and Mazur 2001; Judson and Sawada 2002; Poulis et al. 1998).

Whether using clickers enhances student learning depends on whether they are used appropriately in conjunction with carefully thought-out questions and class discussions

(Guess 2008). In a Journalism course, students had the least favourable attitudes towards the use of the clickers, but the professor was relatively new to using clickers and primarily used them to take attendance. Discussion was neither used nor encouraged and an average of 1–3 clicker questions were used per class (Duncan 2008).

Using SRS has been found to lead to increased student collaboration (Brewer, 2004; Burnstein and Lederman 2001; Robertson 2000; Wieman and Perkins 2005; Wood 2004), attentiveness (Burnstein and Lederman 2001; Carnevale 2005; Roberts 2005; Steinert and Snell 1999), participation (Wampler 2006), interactivity (Homme et al. 2004) and cooperation (Skiba 2006). Skiba (2006) reported that the use of SRS encouraged active learning and student–faculty interaction. Using SRS can enhance communication in the classroom, help students to become more committed to their learning, and make students more accountable (Wieman and Perkins 2005). Use of SRS have been found to be beneficial for formative assessments (Burnstein and Lederman 2001; Carnevale 2005; Duncan 2006; Hatch, Jensen and Moore 2005; Homme et al. 2004; Lightstone 2006; Roberts 2005; Wieman and Perkins 2005; Wood 2004) and for increasing student engagement (Julian 1995; Lightstone 2006; Wood 2004).

Learning environments

Much progress has been made since the late 1960s when Herbert Walberg and Rudolf Moos began their semi-independent research programs on classroom climates (Fraser 2012, 2014). Harvard Project Physics involved a set of research and evaluation activities that led to the creation of the Learning Environment Inventory (LEI) (Walberg and Anderson 1968). During this time, Moos developed the first scales that measured the social climates of psychiatric hospitals and correctional institutions (Moos 1974), which eventually led to the creation of the Classroom Environment Scale (CES) (Moos 1979; Moos and Trickett 1974).

The ideas of Lewin (1936) and Murray (1938) and their followers have often guided research on learning environments. In 1936, Lewin realised the importance of the environment as well as its interaction with individuals. Even with such an abundance of work being focused on learning environments, Fraser (2001) explains how teachers often speak about classroom climate, but very seldom effectively evaluate it.

Fraser (1998, p. 527) stated that "although research and evaluation in science education have relied heavily on the assessment of academic achievement and other valued learning outcomes, these measures cannot give a complete picture of the educational process". That is, too often we rely solely on student performance on tests to evaluate what is happening in a classroom and to evaluate teacher effectiveness and student progress. Fraser (1998, p. 528) also claims that "students are at a good vantage point to make judgments about classrooms because they have encountered many different learning environments and have enough time in a class to form accurate impressions. Also, even if teachers are inconsistent in their day-to-day behavior, they usually project a consistent image of the long-standing attributes of classroom environment."

Throughout the years, as research has shown the importance of effective tools for measuring classroom climate, numerous instruments have been created to do just that. Many of these instruments measure specific scales based on Moos' (1974) scheme for classifying dimensions of human environments: relationship dimensions identify the nature and intensity of personal relationships within the environment and assesses the extent to which people are involved in the environment and support and help each other; personal development dimensions assess the basic directions along which personal growth and self-enhancement tend to occur; and system maintenance and system change dimensions involve the extent to which the environment is orderly, clear in expectations, maintains control and is responsive to change. These instruments are used worldwide and much research is still being conducted to cross-validate them in different languages (Fraser 2012, 2014).

As noted above, the LEI and CES are historically-significant instruments. The LEI contains seven statements in each of 15 different scales with respondents stating whether they agree or disagree with the statement using the responses of Strongly Agree, Agree, Disagree, and Strongly Disagree (Fraser, Anderson and Walberg 1982; Walberg and Anderson 1968). The CES has 10 items in nine different scales (Involvement, Affiliation, Teacher Support, Task Orientation, Competition, Order and Organization, Rule Clarity, Teacher Control, and Innovation) with a True–False response format (Fisher and Fraser 1983; Moos 1979; Moos and Trickett 1974). The My Class Inventory (MCI) is a simplified version of the LEI for children aged 8–12 years (Fisher and Fraser 1981) with 6–9 items in the following scales: Cohesiveness, Friction, Satisfaction, Difficulty, and Competitiveness (Fisher and Fraser 1981; Fraser et al. 1982; Fraser and O'Brien 1985; Goh, Young and Fraser 1995.

Because of the uniqueness of science laboratory classes, the Science Laboratory Environment Inventory (SLEI) was developed with seven items in each of five scales (Student Cohesiveness, Open-Endedness, Integration, Rule Clarity, and Material Environment) and with frequency response alternatives of Almost Never, Seldom, Sometimes, Often, and Very Often. This questionnaire has been validated in numerous countries (Fisher, Henderson and Fraser 1997; Fraser, Giddings and McRobbie 1995; Fraser and Lee 2009; Fraser and McRobbie 1995; Lightburn and Fraser 2007; Wong and Fraser 1995).

The Constructivist Learning Environments Survey (CLES) assesses the degree to which a classroom's environment is consistent with a constructivist epistemology using the scales of Personal Relevance, Uncertainty Critical Voice, Shared Control, and Student Negotiation. The CLES has been cross-validated in various countries (Aldridge, Fraser and

Sebela 2004; Aldridge, Fraser, Taylor and Chen 2000; Kim, Fisher and Fraser 1999; Koh and Fraser 2014; Nix, Fraser and Ledbetter 2005; Peiro and Fraser 2009; Taylor, Fraser and Fisher 1997). The What Is Happening In this Class? (WIHIC) questionnaire combines scales from previous instruments and incorporates them with some new scales (Fraser, Fisher and McRobbie 1996). The WIHIC is the most-frequently used learning environment instrument around the world today and has been found to be valid and useful in studies in Australia and Taiwan (Aldridge and Fraser 2000), Australia, the UK and Canada (Dorman, 2003), Australia and Indonesia (Fraser, Aldridge and Adolphe 2010), the USA (Helding and Fraser 2013; Taylor and Fraser 2013; Wolf and Fraser 2008), Canada (Fraser and Raaflaub 2013), Singapore (Chionh and Fraser 2009; Khoo and Fraser 2008; Peer and Fraser in press), the United Arab Emirates (Afari et al. 2013; MacLeod and Fraser 2010) and South Africa (Aldridge, Fraser and Ntuli 2009). Because of the relevance of what the WIHIC measures, numerous scales (Involvement, Task Orientation, Equity, and Cooperation) from it were used when developing the How Do You Feel About This Class? questionnaire for the present study.

Student attitudes

Baron and Byrne (1977) describe attitudes as individually-attributed beliefs, emotions and behavioural tendencies that someone has towards specific abstract or concrete objects. Attitudes are major determinants of behavior (Tavsancil 2006). Krech and Crutchfield (1980) explain that understanding attitudes allows knowledge of several related behaviours. Emotions, which are expressed through attitudes, affect what is being learned and have a significant impact on learning (Caine and Caine 1994; Lackney 1998). According to Stodolsky, Salk and Blaessner (1991), even if information is forgotten, attitudes towards a subject often remain. Allport (1956) reported that the first study of attitudes was conducted by Thurstone (1929).

Compared to most other subjects taught in schools, science seems to be the one in which researchers have invested the most time in investigating attitudes (Tytler and Osborne 2012). These attitudes in science can describe students' enjoyment of a science class, enjoyment in manipulating equipment (such as in a laboratory setting), enjoyment of their pursuit of knowledge, and interest in pursuing a career in science (Wolf and Fraser 2008). Students' attitudes towards their middle-school science classes can have a major impact on their choice of science courses in high school and college (Misiti, Shrigley and Hanson 1991).

Because the term 'attitude' can take on many meanings, Klopfer (1971, 1976) identified six distinct categories of conceptually-different attitudinal aims: manifestation of favourable attitude towards science and scientist, acceptance of scientific inquiry as a way of thought, adoption of scientific attitudes, enjoyment of science learning experiences, development of interest in science and science related activities, and development of interest in pursuing a career in science. Our study involved assessing attitudes that are primarily linked to Klopfer's categories of manifestation, adoption and enjoyment.

Perrodin (1966) assessed attitudes among a sample of over 500 grade 4, 6, and 8 students in the USA through the use of open-ended statements which allow students to input their own feedback. This qualitative method required a great amount of time in collecting, transcribing and analysing data.

To assess emotional and intellectual attitudes towards science among secondary school students, Moore and Sutman (1970) developed the Scientific Attitude Inventory (SAI), which contains 60 items that range from the knowledge of laws and theories of science to feelings about being a scientist. After examining 30 studies that used the SAI, Munby (1983) questioned its validity. It was conceptualised by Baker (1985) that the SAI possesses two scales, positive and negative. Because of this, he calculated the total attitude score as the score on the positive scale minus the score on the negative scale. Other studies did not report such separation of scales or calculations (Munby 1982). The SAI was revised by Moore and Foy (1997), who did not discuss any changes to overcome these difficulties, which allows for the continuation of doubts as to its validity and reliability (Munby 1997).

Because Fraser (1978) noted three potential problems with several existing instruments used to assess attitudes towards science (low statistical reliability, a lack of economy of items, and the combination of distinct attitude concepts into a single scale which creates a mixture of variables), he developed the Test of Science Related Attitudes (TOSRA) to overcome these problems. The TOSRA is a multidimensional questionnaire with different scales to assess the conceptually-distinct attitude constructs identified by Klopfer (1971). These scales were extended and improved in various ways and two new scales were also added to create the final version of the TOSRA with seven scales consisting of 10 items each (Fraser, 1981). Not only has TOSRA been found to be valid and useful in science classes in numerous countries such as Singapore (Wong and Fraser 1996), Australia and Indonesia (Fraser, Aldridge and Adolphe 2010) and Korea (Fraser and Lee 2009), but it has been modified and cross-validated for other school subjects, including mathematics (Ogbuehi and Fraser 2007), geography (Walker 2006) and English (Liu and Fraser 2013). Because of the relevance of the TOSRA to our study, one scale (Enjoyment of Science Lessons) was used in the creation of the How Do You Feel About This Class? questionnaire. The original 10-item scale from TOSRA was reduced to 8 items (for economy) and any negatively-worded items were converted to positive wording (to avoid confusing students).

Methods

This study used a new questionnaire, the How Do You Feel About This Class? (HDYFATC), which incorporates numerous scales from the WIHIC, one scale from the TOSRA, and a scale (called Comfort) created by the researchers. This new questionnaire was used to assess students' views of their learning environment and their attitudes. Table 1 gives a description and sample item for each scale. The response alterations for HDYFATC items were Strongly Disagree, Disagree, Not Sure, Agree and Strongly Agree. Students' achievement test scores for the duration of the study provided information for determining the effectiveness of Student Response Systems in terms of achievement. The achievement scores were based on teachers' normal quizzes and examinations. The use of the same assessments allowed for consistency. At the end of the study, an average score was determined based in each quiz and examination grade. These averages were then divided by 20 for consistency with the range of scores possible for HDYFATC scales.

Table 1 here

Our evaluation of the use of SRS involved comparing two instructional groups (SRS and non-SRS). In an attempt to minimise the teacher as a variable in the study, each teacher involved used SRS in half of his/her classes but not in the other half. The amount of time for which SRS was used was approximately four months.

Data were collected from 1,097 grade 7 and 8 science students from 47 classes in southern New York State. Of these students, 544 were male and 553 were female. These students ranged from advanced, to average, to special education. The SRS group consisted of 532 students (266 males and 266 females). The 565 students in the non-SRS group was composed of 544 males and 553 females.

Student responses to the questionnaire were used to cross-validate the questionnaire using principal axis factoring followed by varimax rotation and Kaiser normalisation. Only items with a factor loading of 0.40 or above on its own scale and less than 0.40 on all other scales were retained. A one-way ANOVA was used to determine the ability of each scale to differentiate between the perceptions of students in different classes. MANOVA was used to compare two groups, those who used SRS and those who did not, in terms of learning environment, attitude and achievement scales. Differences between the two groups also were described in terms of the effect size (magnitude of the differences in standard deviations) and the statistical significance for each scale.

Results

Validity and reliability of the HDYFATC

For a sample of 1,097 students, we checked the structure of the HDYFATC using principal axis factor analysis with varimax rotation and Kaiser normalisation. In order for an item to be retained, it needed a factor loading of at least 0.40 on its a priori scale and less than 0.40 on every other scale. Factor loadings for all items of the HDYFATC were above 0.40 on their a priori scale, ranging from 0.43 to 0.79, and no item had a loading greater than 0.40 on a different scale. Therefore all 48 items and all six scales were retained. The factor analysis results are shown in Table 2.

Table 2 here

The percentage of the total variance extracted with each factor ranged from 2.70% to 53.38% for different scales, with the total variance accounted for being 76.13%. The largest contribution to variance was for Enjoyment (53.38%). Eigenvalues ranged from 1.29 to 25.62 for different scales. The results of the factor analysis strongly support the factorial validity of the final 48-item, six-scale version of the HDYFATC when used with our sample of middle-school students in New York.

For each of the six scales of the HDYFATC, the internal consistency reliability was estimated for two units of analysis (the student and the class mean), using the Cronbach alpha coefficient. Table 2 shows that, when the individual was used as the unit of analysis, the alpha coefficient for different scales ranged from 0.94 to 0.95. With the class mean as the unit of analysis, the internal consistency reliability ranged from 0.97 to 0.99 for different scales. Reliability estimates were higher when the class mean was used as the unit of analysis.

Through using an ANOVA, the ability of each learning environment scale of the HDYFATC to differentiate between perceptions of students in different classrooms was determined. ANOVA indicates if students in the same class perceive their learning environment in a similar way, while mean class perceptions vary from class to class. The results reported in Table 2 reveal a significant difference between students' perceptions in different classes for each learning environment scale of the HDYFATC. The eta^2 statistic, which represents the proportion of variance in scale scores accounted for by class membership, ranged from 0.50 to 0.60 for the different learning environment scales. (This characteristic is not relevant for the Enjoyment scale.)

As in considerable past research (Aldridge and Fraser 2000; Chionh and Fraser 2009; den Brok, Fisher, Rickards and Bull 2006; MacLeod and Fraser 2010; Martin-Dunlop and Fraser 2008; Wolf and Fraser 2008), scales from the WIHIC and TOSRA showed strong validity and reliability.

Effectiveness of SRS

To determine the effectiveness of the use of SRS in terms of learning environment, attitudes, and achievement, each scale's average item mean (the scale mean divided by the number of items in a scale) and average item standard deviation were calculated (Table 3). As recommended by Thompson (1998, 2002), effect sizes were also calculated to describe the magnitude of the difference between the SRS and control groups. Effect sizes (Cohen's *d*) show the differences between means expressed in standard deviation units (the difference between the means of two groups divided by the pooled standard deviation). These results for our study are shown in Table 3.

To ascertain the statistical significance of differences between the two instructional groups, a one-way multivariate analysis of variance (MANOVA) was conducted with the five learning environment scales and student outcome scales (achievement and enjoyment) as the dependent variables and the use or non-use of SRS as the independent variable. Because the multivariate test using Wilks' lambda criteria yielded a statistically significant result overall for the whole set of seven dependent variables, the univariate ANOVA results were interpreted separately for each individual dependent variable. Table 3 provides ANOVA results, as well as the effect size, for each of the seven learning environment and student outcome variables.

Table 3 shows that students in classes in which SRS were used had statistically significantly higher scores on all learning environment scales contained in the HDYFATC than did students in the control group. Effect sizes ranged from 1.96 to 2.46 standard deviations. Also, students who used SRS enjoyed science classes statistically significantly more than students who did not use SRS, with an effect size of 2.19 standard deviations, and had statistically significantly higher achievement, with an effect size of 1.17 standard deviations. These effect sizes are remarkably large according to Cohen's (1992) criteria.

Table 3 here

Previous studies of the effectiveness of SRS have revealed similar results in terms of increased involvement (Duncan 2008; Joosten and Kaleta 2006; Lightstone 2006; Martyn 2007; Wampler 2006), increased equity (Duncan 2008; Martyn 2007), students feeling more comfortable in their science class (Duncan 2008; Martyn 2007), more favourable attitudes (Duncan 2008; Martyn 2007; Roberts 2005; Siau et al. 2006), increased student achievement (Guess 2008; Martyn 2007), and an increase in task orientation and cooperation (Guess 2008; Skiba 2006).

Significance and Implications

This evaluation of SRS at the middle-school level (grade 7 and 8) is distinctive because most past research on the effectiveness of SRS has been conducted at the higher levels of education. The questionnaire used in our study, How Do You Feel About This Class? (HDYFATC), was shown to have sound factorial validity and internal consistency reliability, as well as being able to differentiate between the perceptions of students in different classrooms, for a sample of 1097 middle-school science students in New York. This questionnaire, which takes only approximately 10 minutes for students to complete, can be used with confidence by future researchers and teachers to assess students' perceptions of the learning environment and their attitudes.

Our study suggests that the use of SRS in science classrooms can help to improve student perceptions of the learning environment, their attitudes towards science, and their achievement. Our carefully-controlled comparison of SRS and non-SRS groups revealed very large differences of 1.17–2.45 standard deviations for seven learning environment, attitude and achievement criteria. School districts can use the findings from our study to help them to decide if investing a portion of their monetary budget on this specific technology is likely to be beneficial to their students. Although many schools attempt to maintain the latest technology when possible, this new technology is very expensive during the current economic crisis. Districts might only wish to invest in technology that has been shown by research to have a positive impact on students.

References

Afari, E., Aldridge, J. M., Fraser, B. J., & Khine, M. S. (2013). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research, 13,* 131-150.

Aldridge, J.M., Fraser, B.J., & Ntuli, S. (2009). Utilising learning environment assessments to improve teaching practices among in-service teachers undertaking a distance education programme. *South African Journal of Education, 29,* 147-170.

Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research, 3*, 101-134.

Aldridge, J.M., Fraser, B.J., & Sebela, M.P. (2004). Using teacher action research to promote constructivist learning environments in South Africa. *South African Journal of Education, 24,* 245-253.

Aldridge, J.M., Fraser, B.J., Taylor, P.C., & Chen, C.-C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22,* 37-55.

Allport, G. (1956). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Baker, D. R. (1985). Predictive value of attitude, cognitive ability, and personality to science achievement in the middle school. *Journal of Research in Science Teaching*, *22*, 103-113.

Baron, R. A., & Byrne, D. (1977). *Understanding human interaction* (2nd ed.). Boston: Allyn and Bacon.

Brewer, C. (2004). Near real-time assessment of student learning and understanding in biology courses. *BioScience*, *54*, 1034-1040.

Burnstein, R., & Lederman, L. (2001). Using wireless keypads in lecture classes. *Physics Teaching, 39*, 8-11.

Caine, G., & Caine, R. N. (1994). *Making connections: Teaching and the human brain*. Menlo Park, CA: Addison-Wesley.

Carnevale, D. (2005). Run class like a game show: 'Clickers' keep students involved. *The Chronicle of Higher Education, 51*, 42.

Casanova, J. (1971). An instructional experiment in organic chemistry: The use of student response systems. *Journal of Chemical Education, 48*, 453-455.

Chionh, Y. H., & Fraser, B. J. (2009). Classroom environment, achievement, attitudes and self-esteem in geography and mathematics in Singapore. *International Research in Geographical and Environmental Education, 18*, 29-44.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*, 970-977.

den Brok, P., Fisher, D., Rickards, T. W., & Bull, E. (2006). Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation*, *12*(1), 3-25.

Dorman, J. P. (2003). Cross-national validation of the What Is Happening In this Class? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research, 6,* 231-245.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, *20*, 81-94.

Duncan, D. (2005). *Clickers in the classroom: How to enhance science teaching using classroom response systems*. San Francisco, CA: Pearson Education Inc.

Duncan, D. (2006). Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review, 5*, 70-88.

Duncan, D. (2008, November). *Clickers: A new teaching tool of exceptional promise.* Paper presented at the Inaugural Conference on Classroom Response Systems: Innovations and Best Practices, University of Louisville, KA.

Fisher, D. L., & Fraser, B. J. (1981). Validity and use of My Class Inventory. *Science Education, 65*, 145-156.

Fisher, D. L., & Fraser, B. J. (1983). Validity and use of Classroom Environment Scale. *Educational Evaluation and Policy Analysis, 5*, 261-271.

Fisher, D. L., Henderson, D., & Fraser, B. J. (1997). Laboratory environments and student outcomes in senior high school biology. *American Biology Teacher, 59,* 214-219.

Fraser, B. J. (1978). Some attitude scales for ninth grade science. *School Science and Mathematics, 78,* 379-384.

Fraser, B. J. (1981). *Test of Science-Related Attitudes handbook*. Melbourne, Victoria: Australian Council for Educational Research.

Fraser, B. J. (1998). Science learning environments: Assessment, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 527-564). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fraser, B. J. (2001). Twenty thousand hours: Editor's introduction. *Learning Environments Research, 4*, 1-5.

Fraser, B. J. (2012). Classroom learning environments: Retrospect, context and prospect. In B. J. Fraser, K. G. Tobin & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191-1240). New York: Springer.

Fraser, B. J. (2014). Classroom learning environments: Historical and contemporary perspectives. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education.* Volume II (pp. 104-117). New York: Routledge.

Fraser, B.J., Aldridge, J.M., & Adolphe, F.S.G. (2010). A cross-national study of secondary science classroom environments in Australia and Indonesia. *Research in Science Education, 40,* 551-571.

Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (3rd ed.). Perth, Australia: Western Australian Institute of Technology.

Fraser, B. J., Fisher, D. L., & McRobbie, C. J. (1996, April). *Development, validation, and use of personal and class forms of a new classroom environment instrument.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching, 32*, 399-422.

Fraser, B. J., & Lee, S. S. U. (2009). Science laboratory classroom environments in Korean high schools. *Learning Environments Research, 12,* 67-84.

Fraser, B. J., & McRobbie, C. J. (1995). Science laboratory classroom environments at schools and universities: A cross-national study. *Educational Research and Evaluation, 1*, 289-317.

Fraser, B. J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary-school classrooms. *Elementary School Journal, 85*, 576-580.

Fraser, B. J. & Raaflaub, C. (2013). Subject and sex differences in the learning environments – Perceptions and attitudes of Canadian mathematics and science students using laptop computers. *Curriculum and Teaching, 28* (1), 57-58.

Goh, S. C., Young, D. J., & Fraser, B. J. (1995). Psychosocial climate and student outcomes in elementary mathematics classrooms: A multilevel analysis. *Journal of Experimental Education, 64*, 29-40.

Guess, A. (2008). *Keeping clickers in the classroom*. Retrieved 25 March 2009, from http://www.insidehighered.com/news/2008/07/18/clickers

Hatch, J., Jensen, M., & Moore, R. (2005). Manna from heaven or 'clickers' from hell. *Journal of College Science Teaching, 34*, 36-39.

Helding, K., & Fraser, B. J. (2013). Effectiveness of National Board Certification (NBC) teachers in terms of classroom environment, attitudes and achievement among secondary science students. *Learning Environments Research, 13*, 1-21.

Homme, J., Asay, G., & Morgenstern, B. (2004). Utilization of an audience response system. *Medical Education, 38*, 575.

Joosten, T., & Kaleta, R. J. (2006, April). *"Clickers" in the classroom: Analyses from the University of Wisconsin System Project.* Paper presented at the UW System Joint Conference 2006, Green Lake, WI.

Judson, E., & Sawada, D. (2002). Learning from past and present: Electronic response systems in college lecture halls. *Journal of Computers in Mathematics and Science Teaching, 21*, 167-181.

Julian, G. M. (1995). Socratic dialogue – With how many? *The Physics Teacher, 33*, 338-339.

Khoo, H. S., & Fraser, B. J. (2008). Using classroom psychosocial environment in the evaluation of adult computer application courses in Singapore. *Technology, Pedagogy and Education, 17,* 67-81.

Kim, H. B., Fisher, D. L., & Fraser, B. J. (1999). Assessment and investigation of constructivist science learning environments in Korea. *Research in Science and Technological Education, 17*, 239-249.

Klopfer, L. E. (1971). Evaluation of learning in science. In B. S. Bloom, J. T. Hastings & G. F. Madaus (Eds.), *Handbook on summative and formative evaluation of student learning* (pp. 559-641). New York: McGraw Hill.

Klopfer, L. E. (1976). A structure for the affective domain in relation to science education. *Science Education, 60*, 299-312.

Koh, N. K., & Fraser, B. J. (2014). Learning environment associated with the use of mixed-mode delivery model among secondary business studies students in Singapore. *Learning Environments Research, 17*, 157-171.

Kretch, D., & Crutchfield, R. S. (1980). *Sosyal psikoloji: Teori ve problemler.* Istanbul: Istanbul Üniversitesi.

Lackney, J. (1998). *Design principles based on brain-based learning research*. Retrieved 8 June, 2010, from http://www.designshare.com/research/brainbasedlearn98.htm

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw.

Lightburn, M. E., & Fraser, B. J. (2007). Classroom environment and student outcomes among students using anthropometry activities in high-school science. *Research in Science and Technological Education, 25*, 153-166.

Lightstone, K. (2006). How remote responders affect teaching. *The Teaching Professor, 20*, 8.

Littauer, R. (1972). Instructional implications of a low-cost electronic student response system. *Educational Technology Teacher and Technology Supplement, 12*, 69-71.

Liu, L., & Fraser, B. J. (2013). Development and validation of an English classroom learning environment inventory and its application in China. In M. S. Khine (Ed.), *Application of structural equation modeling in educational research* (pp. 75-89). Rotterdam: Sense Publishers.

MacLeod, C., & Fraser, B. J. (2010). Development, validation and application of a modified Arabic translation of the What Is Happening In this Class? (WIHIC) questionnaire. *Learning Environments Research, 13*, 105-125.

Martin-Dunlop, C., & Fraser, B. J. (2008). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education, 6*, 163-190.

Martyn, M. (2007). Clickers in the classroom: An active learning approach. *Educase Quarterly, 30*(2), 71-74.

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.

Misiti, F. L., Shrigley, R. L., & Hanson, L. (1991). Science attitudes scale for middle school students. *Science Education, 75*, 525-540.

Moore, R. W., & Foy, R. L. (1997). The Scientific Attitude Inventory: A revision (SAI II). *Journal of Research in Science Teaching, 34*, 327-336.

Moore, R. W., & Sutman, F. X. (1970). The development, field test and validation of an inventory of scientific attitudes. *Journal of Research in Science Teaching, 34*, 327-336.

Moos, R. H. (1974). *The social climate scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.

Moos, R. H. (1979). *Evaluating educational environments: Procedures, measures, findings and policy implications*. San Francisco: Jossey-Bass.

Moos, R. H., & Trickett, E. J. (1974). *Classroom Environment Scale manual*. Palo, Alto, CA: Consulting Psychologists Press.

Munby, H. (1982). The impropriety of "panel of judges" validation in science attitude scales: A research comment. *Journal of Research in Science Teaching, 19,* 617-619.

Munby, H. (1983). Thirty studies involving the "Scientific Attitude Inventory": What confidence can we have in this instrument? *Journal of Research in Science Teaching, 20*, 141-162.

Munby, H. (1997). Issues in validity of science attitude measures. *Journal of Research in Science Teaching, 34*, 337-341.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Nix, R. K., Fraser, B. J., & Ledbetter, C. E. (2005). Evaluating an integrated science learning environment using the Constructivist Learning Environment Survey. *Learning Environments Research, 8*, 109-133.

Ogbuehi, P. I., & Fraser, B. J. (2007). Learning environment, attitudes and conceptual development associated with innovative strategies in middle-school mathematics. *Learning Environments Research, 10*, 101-114.

Peer, J., & Fraser, B. J. (in press). Sex, grade-level and stream differences in learning environment and attitudes to science in Singapore primary schools. *Learning Environments Research.*

Peiro, M. M., & Fraser, B. J. (2009). Assessment and investigation of science learning environments in the early childhood grades. In M. Ortiz & C. Rubio (Eds.), *Educational evaluation: 21st century issues and challenges* (pp. 349-365). New York: Nova Science Publishers.

Perrodin, A. F. (1966). Children's attitudes towards science. *Science Education, 50*, 214-218.

Poulis, J., Massen, C., Robens, E., & Dilbert, M. (1998). Physics lecturing with audience paced feedback. *American Journal of Physics, 66*, 439-441.

Roberts, G. (2005). Instructional technology that's hip high-tech. *Computer in Libraries, 25*, 26-28.

Robertson, L. J. (2000). Twelve tips for using a computerized interactive audience response system. *Medical Teaching, 22*, 237-239.

Siau, K., Sheng, H., & Nah, F. F.-H. (2006). Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education, 49*, 399-403.

Skiba, D. (2006). Got large lecture halls? Use clickers. *Nursing Education Perspective, 27*, 278-280.

Steinert, Y., & Snell, L. S. (1999). Interactive lecturing: Strategies for increasing participation in large group presentations. *Medical Teaching, 21*, 37-42.

Stodolsky, S. S., Salk, S., & Blaessner, B. (1991). Student views about learning math and social sciences. *American Educational Research Journal, 28*, 89-116.

Tavsancil, E. (2006). *Tutumlarin Ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Yayinlari.

Taylor, B. A., & Fraser, B. J. (2013). Relationships between learning environment and mathematics anxiety. *Learning Environments Research, 16*, 297-313.

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293-302.

Thompson, B. (1998). Review of 'What if there were no significance tests?' *Educational and Psychological Measurement, 58*, 334-346.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher, 31*, 24-31.

Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Bulletin, 36*, 222-241.

Tytler, R., & Osborne J. (2012). Student attitudes and aspirations towards science. In B. J. Fraser, K. G. Tobin & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 597-625). New York: Springer.

Walberg, H. J., & Anderson, G. J. (1968). Classroom learning and individual learning. *Journal of Educational Psychology, 59*, 414-419.

Walker, S. L. (2006). Development and validation of the Test of Geography-Related Attitudes (ToGRA). *Journal of Geography, 105*, 175-181.

Wampler, P. J. (2006). Clickers in the classroom – Rewards and regrets of using student response systems in a large enrollment geology course. *Geological Society of America Abstracts with Programs, 38*, 497.

Wieman, C., & Perkins, K. (2005). Transforming physics education. *Physics Today*, 36-41.

Wit, E. (2003). Who wants to be... The use of a personal response system in statistics teaching. *MSOR Connections, 3*(2), 14-20.

Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education, 38*, 321-341.

Wong, A. F. L., & Fraser, B. J. (1996). Environmental−attitude associations in the chemistry laboratory classroom. *Research in Science and Technological Education, 14*, 91-102.

Wood, W. B. (2004). Clickers: A teaching gimmick that works. *Developmental Cell, 7*, 796-798.

Table 1 Scale descriptions and sample item for each HDYFATC scale

Scale	Scale description	Sample item
Involvement	Extent to which students participate in discussions, perform additional work, and are attentive during the class.	I give my opinions during class discussions.
Task Orientation	Extent to which students feel that it is important to complete activities and to stay on task.	Getting a certain amount of work done is important to me.
Cooperation	Extent to which students cooperate with each other rather than compete when completing tasks.	Students work with me to achieve class goals.
Equity	Extent to which students feel as though they are treated equally by the teacher.	The teacher gives as much attention to my questions as to other students' questions.
Comfort	Extent to which students feel comfortable and safe in participating in class discussions and answer questions posed to the class.	I am comfortable when raising my hand to participate in this class.
Enjoyment	Extent to which students enjoy their class and look forward to going to it.	I enjoy going to science class.

The response alternatives used for the HDYFATC were Strongly Disagree, Disagree, Not Sure, Agree and Strongly Agree.

Table 2 Factor analysis results, internal consistency (Cronbach alpha reliability) and ability to differentiate between classrooms (ANOVA results) for HDYFATC

Item No	Factor loadings					
	Enjoyment	Involvement	Task Orientation	Cooperation	Equity	Comfort
1	0.60					
2	0.73					
3	0.76					
4	0.65					
5	0.63					
6	0.62					
7	0.65					
8	0.79					
9		0.66				
10		0.75				
11		0.68				
12		0.66				
13		0.63				
14		0.70				
15		0.66				
16		0.70				
17			0.74			
18			0.57			
19			0.65			
20			0.72			
21			0.66			
22			0.61			
23			0.63			
24			0.77			
25				0.43		
26				0.71		
27				0.72		
28				0.50		
30				0.56		
31				0.70		
32				0.53		
33					0.72	
34					0.68	
35					0.58	
36					0.65	
37					0.66	
3					0.65	
39					0.73	
40					0.69	
41						0.59
42						0.65
43						0.72
44						0.63
45						0.76
46						0.74
47						0.73
48						0.78
% Variance	53.38	6.60	5.39	4.61	3.45	2.70
Eigenvalue	25.62	3.17	2.58	2.21	1.65	1.29
ANOVA (eta^2)	a	0.59*	0.52*	0.50*	0.55*	0.60*
Alpha Reliability						
Student	0.95	0.95	0.95	0.94	0.94	0.95
Class	0.99	0.98	0.97	0.98	0.97	0.99

N = 1097 students in 47 Classes. *$p<0.001$
Factor loadings less than 0.40 have been omitted from the table.
Principal axis factoring with varimax rotation and Kaiser normalization.
The eta² statistic (which is the ration of “between” to “total” sums of squares) represents the proportion of variance explained by class membership.
[a] Not relevant to the Enjoyment scale

Table 3 Average item mean, average item standard deviation and difference between SRS and control groups (effect size and ANOVA result) for students' perceptions of learning environment, enjoyment and achievement

Scale	Average item mean		Average item SD		Difference
	SRS	Control	SRS	Control	Effect size & significance
Learning Environment					
Involvement	3.91	1.93	0.83	0.81	2.45*
Task Orientation	3.87	2.08	0.76	1.01	2.00*
Cooperation	3.86	2.09	0.94	0.86	1.96*
Equity	3.88	2.06	0.86	0.81	2.17*
Comfort	4.07	2.06	0.77	0.86	2.46*
Student Outcomes					
Enjoyment	3.88	2.00	0.82	0.89	2.19*
Achievement	3.52	3.18	0.29	0.28	1.17*

*$p < 0.001$

Sample consists of 532 students in SRS group and 565 students in control group.

Achievement scores were divided by 20 to be consistent with the score range for questionnaire items.