An Investigation of Social Responsibility: The SRQS

by

Sophie Vararatana Ty

B.A., McGill University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The Faculty of Graduate Studies

(School Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2013

© Sophie Vararatana Ty, 2013
Abstract

This study presents the first validation data for a brief measure of student social responsibility, the Social Responsibility Quick Scale (SRQS). Teachers from two schools from the same school district implementing school-wide positive behaviour support (SWPBS) completed the SRQS for 360 students (51%) at the beginning (September 2011) and end of the school year (June 2012). Content validity procedures provided evidence that the items on the SRQS measure behaviours associated with student social responsibility. Factor analytic procedures demonstrated that the SRQS possesses a one-factor structure. Reliability procedures demonstrated strong internal consistency and test-retest reliability, as well as strong inter-item and corrected item-total correlations. Teacher ratings were consistently and significantly higher for female versus male students. No significant differences were found across grade level. Results are discussed in terms of implications for practice and future research.
Preface

The present study was conducted by the graduate student, under the supervision and direction of her research supervisor. The research supervisor was responsible for the recruitment of the participating schools. Data collection was conducted via a contact at each school, under the supervision of the graduate student and guidance of her research supervisor. The graduate student was responsible for the analysis and writing of the present study. A member of the graduate student’s thesis committee provided advisement during the analytic process. Taken together, the present thesis is representative of the graduate student’s work, namely as the primary researcher and lead author. Approval to conduct research in the school district and ethics approval from the UBC Behavioural Research Ethics Board (BREB) were sought and granted prior to conducting the study. The UBC BREB certificate number is H11-01356.
# Table of Contents

Abstract........................................................................................................................................ ii
Preface........................................................................................................................................... iii
Table of Contents........................................................................................................................... iv
List of Tables ..................................................................................................................................... vi
List of Figures ................................................................................................................................. vii
Acknowledgements ....................................................................................................................... viii
Dedication ......................................................................................................................................... ix

Chapter 1: Introduction .................................................................................................................. 1
  1.1 Social-Emotional Competence ............................................................................................... 2
  1.2 Social Responsibility ............................................................................................................... 3
  1.3 Universal Strength-Based Screening...................................................................................... 5
      1.3.1 From a Traditional to Multi-Tiered Approach .............................................................. 5
      1.3.2 From a Deficit- to a Strength-Based Approach ........................................................... 6
  1.4 Measuring Social Responsibility with Validity and Reliability ........................................... 7
  1.5 Strength-Based Screening Instruments ............................................................................... 8
  1.6 Purpose of the Study ............................................................................................................. 9

Chapter 2: Method ........................................................................................................................ 11
  2.1 Participants and Settings ...................................................................................................... 11
  2.2 Measures ............................................................................................................................. 12
      2.2.1 Social Responsibility Quick Scale (SRQS) ............................................................... 12
      2.2.2 Content Validity Survey ............................................................................................ 13
      2.2.3 Office Discipline Referrals (ODRs) .......................................................................... 13
      2.2.4 Suspensions ............................................................................................................... 14
  2.3 Procedures ............................................................................................................................. 14
      2.3.1 Content Validity ........................................................................................................ 14
      2.3.2 Recruitment of Sites................................................................................................... 15
      2.3.3 Data Collection ......................................................................................................... 15
  2.4 Design and Analyses .......................................................................................................... 15
      2.4.1 Content Validity ........................................................................................................ 16
      2.4.2 Factor Structure ....................................................................................................... 17
      2.4.3 Reliability ................................................................................................................ 18
      2.4.4 Construct Validity ................................................................................................... 18

Chapter 3: Results ......................................................................................................................... 21
  3.1 Content Validity .................................................................................................................... 21
      3.1.1 Inter-Rater Agreement (IRA) ..................................................................................... 21
      3.1.2 Content Validity Index (CVI) .................................................................................. 21
  3.2 Factor Structure .................................................................................................................... 22
      3.2.1 Principal Component Analysis (PCA) ...................................................................... 22
  3.3 Reliability .............................................................................................................................. 23
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.1 Internal Consistency Reliability</td>
<td>23</td>
</tr>
<tr>
<td>3.3.2 Inter-Item and Corrected Item-Total Correlation</td>
<td>24</td>
</tr>
<tr>
<td>3.3.3 Test-Retest Reliability</td>
<td>24</td>
</tr>
<tr>
<td>3.4 Construct Validity</td>
<td>25</td>
</tr>
<tr>
<td>3.4.1 Teacher SRQS Ratings and ODR Data</td>
<td>25</td>
</tr>
<tr>
<td>3.4.2 Teacher SRQS Ratings and Suspension Data</td>
<td>25</td>
</tr>
<tr>
<td>3.4.3 Teacher SRQS Ratings and Student Gender</td>
<td>25</td>
</tr>
<tr>
<td>3.4.4 Teacher SRQS Ratings and Student Grade Level</td>
<td>25</td>
</tr>
<tr>
<td>Chapter 4: Discussion</td>
<td>26</td>
</tr>
<tr>
<td>4.1 Content Validity</td>
<td>27</td>
</tr>
<tr>
<td>4.2 Factor Structure</td>
<td>29</td>
</tr>
<tr>
<td>4.3 Reliability</td>
<td>29</td>
</tr>
<tr>
<td>4.4 Construct Validity</td>
<td>29</td>
</tr>
<tr>
<td>4.6 Limitations</td>
<td>31</td>
</tr>
<tr>
<td>4.7 Implications for Practice and Future Research</td>
<td>32</td>
</tr>
<tr>
<td>References</td>
<td>34</td>
</tr>
<tr>
<td>Appendices</td>
<td>40</td>
</tr>
<tr>
<td>Appendix A: Social Responsibility Quick Scale (SRQS)</td>
<td>40</td>
</tr>
<tr>
<td>Appendix B: Content Validity Survey</td>
<td>44</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1 Core Social-Emotional Competencies Identified by CASEL ............................................. 2
Table 1.2 Social Responsibility Aspects: Grades 4 to 5 ........................................................................ 4
Table 3.1 Items from the Content Validity Survey as Rated by Experts .............................................. 22
Table 3.2 Factor Loadings for the SRQS (N = 360) ............................................................................... 23
Table 3.3 Inter-Item and Corrected Item-Total Correlations for the SRQS: Coefficient alpha .......... 24
List of Figures

Figure 3.1   Scree Plots for September 2011 and June 2011 .............................................. 23
Figure 3.2   Teacher SRQS Ratings and Grade Level .......................................................... 26
Acknowledgements

It is with deepest gratitude that I acknowledge my research supervisor, Dr. Kent McIntosh, Ph.D. Thank you for the opportunities you made available to me throughout my degree, and the guidance and support you provided throughout my thesis experience. Your knowledge, work ethic, and professionalism have been a source of learning, admiration, and respect. I would also like to extend a thank you to my supportive and knowledgeable thesis committee members, Dr. Ruth Ervin, Ph.D. and Dr. Bruno Zumbo, Ph.D.

Additionally, I would like to acknowledge my SUPPORT Lab peers—Christina Moniz, Sara Greflund, and Mary Turri—as well as the Capsule Crew. Thank you all for being an invaluable source of support and laughter throughout my degree and thesis experience. I have learned so much from each of you in ways that I cherish deeply.

Lastly, I would like to thank my immediate and extended family for their invaluable love and support.
Dedication

I dedicate this thesis to the love of my life, Bruce Ramsay.
Chapter 1: Introduction

Schools play a central role in the lives of children and youth, and although schooling has largely been synonymous with academic instruction, there is increasing recognition of the benefits of integrating social-emotional learning into the educational curriculum (Elias et al., 1997). For example, there is growing acknowledgement by educators and related professionals that a “comprehensive mission of schools is to educate students to be knowledgeable, responsible, socially skilled, healthy, caring, and contributing citizens” (Greenberg et al., 2003, p. 466).

Social-emotional learning and academic achievement are not mutually exclusive, but rather interrelated domains (Zins & Elias, 2007). A growing body of research suggests that a comprehensive instructional approach that includes both domains provides optimal preparation for student success in school and in life (Elias et al., 1997). Results from two meta-analyses found that implementation of school-based social-emotional instruction was associated not only with valued student outcomes such as increased social competence and prosocial behaviour, but also with gains in academic achievement comparable to results from interventions with a strictly educational focus (Durlak, Weissburg, Dymnicki, Taylor, & Schellinger, 2011; Payton et al., 2008). Although much of the data supporting social-emotional instruction are based on U.S. studies, a recent meta-analysis that included a considerable number of studies based in other regions of the world found similar results, lending support to the value of school-based social-emotional instruction for children and youth around the world (Fundación Marcelino Botín, 2008).

Since the establishment of the Collaborative for Academic, Social, and Emotional Learning (CASEL; www.casel.org) in 1994, a U.S.-based organization that has been leading national and international efforts to make social-emotional instruction an integral part of the
educational curriculum for children and youth, the recognition of the importance of social-emotional learning has gained worldwide momentum. For example, in 2003 the International Academy of Education (IAE) and UNESCO’s International Bureau of Education (IBE) created a readily available resource outlining 10 evidence-based best-practice guidelines for social-emotional instruction in schools (Elias, 2003). In Canada, the past decade has seen an emergence of social-emotional awareness initiatives (e.g., *BC Performance Standards, Social Responsibility: A Framework; Character Development in Ontario Schools*).

### 1.1 Social-Emotional Competence

A key outcome of social-emotional learning is social-emotional competence. Generally defined, social-emotional competence is the “ability to understand, manage, and express the social and emotional aspects of one’s life in ways that enable the successful management of life tasks such as learning, forming relationships, solving everyday problems, and adapting to the complex demands of growth and development” (Elias et al., 1997, p.2). CASEL identified five core competencies that make up social-emotional competence (see Table 1.1; www.casel.org). Based on the outline provided by CASEL, it becomes evident that social-emotional competence encompasses a variety of similar concepts and skills.

**Table 1.1 Core Social-Emotional Competencies Identified by CASEL**

<table>
<thead>
<tr>
<th>Social-Emotional Competency</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-management</td>
<td>Managing one’s emotions and behaviours in a way that is conducive to achieving one’s goals</td>
</tr>
<tr>
<td>Self-awareness</td>
<td>Recognizing one’s emotions, values, strengths, and limitations</td>
</tr>
<tr>
<td>Responsible decision-making</td>
<td>Making decisions based on consideration of ethical and social norms, and likely consequences and contributions</td>
</tr>
<tr>
<td>Relationship skills</td>
<td>Initiating and maintaining positive relationships, dealing with conflict effectively, cooperating</td>
</tr>
<tr>
<td>Social awareness</td>
<td>Being able to appreciate other perspectives, show empathy, and recognize available resources</td>
</tr>
</tbody>
</table>

Source: [www.casel.org](http://www.casel.org)
1.2 Social Responsibility

Social responsibility, or socially responsible behaviour with others, is one concept that permeates each of the social-emotional competencies identified by CASEL. The construct of social responsibility encompasses many behaviours, including cooperating, showing respect for others, conforming to social rules and norms, initiating and maintaining positive social interactions, and developing a sense of moral responsibility (Wentzel, 1991). A succinct and apt definition of social responsibility is “adherence to social rules and role expectations” (Wentzel, p. 2). This definition acknowledges that the underlying structure of human relationships is governed by a system of rules and role expectations, which provides the guidelines for the participatory nature of each person in any context in relation to another (Wentzel). In the school context, such a system defines the role of the student. One example of a school-based system of rules and role expectations is school-wide positive behaviour support (SWPBS), a multi-tiered prevention and intervention framework that has as a key practice the explicit teaching, monitoring, and acknowledgement of prosocial behaviours (e.g., raise hand before speaking, help others in need; Horner, Sugai, Todd, & Lewis-Palmer, 2005). Results from randomized controlled effectiveness trials, and quasi-experimental and comparison studies have documented significant effects of SWPBS on problem behaviour (e.g., talking out of turn, distracting peers during lessons) and academic achievement (e.g., province/state and standardized achievement tests; Bradshaw, Mitchell, & Leaf, 2010; Bradshaw, Waasdorp, & Leaf, 2012; Horner et al., 2009; Lassen, Steele, & Sailor, 2006; Luiselli, Putnam, Handler, & Feinberg, 2005; McIntosh, Bennett, & Price, 2011; Muscott, Mann, & Lebrun, 2008; Nelson, Martella, & Marchand-Martella, 2002).

In 2001, the British Columbia Ministry of Education identified social responsibility as a key learning area, as important to education as reading, writing, and numeracy. The BC Ministry
of Education created a document titled, *BC Performance Standards, Social Responsibility: A Framework* to establish a common language, set of expectations, and standards with regard to student social responsibility in BC schools (BC Ministry of Education, 2001). The stated objective of the *Standards* is to promote a sense of social mindedness and social responsibility in children and youth in BC schools (BC Ministry of Education). Extensive field research informed the identification and operationalization of four aspects of social responsibility at various grade levels (see Table 1.2; BC Ministry of Education). For each aspect of social responsibility, the *Standards* provides a description of observable behaviours for each of four different levels of performance expectations. For example, a Grade 4 student who “fully meets expectations” with regard to “contributing to the classroom and school community” will display the key skills as described in Table 1.2.

Table 1.2  Social Responsibility Aspects: Grades 4 to 5

<table>
<thead>
<tr>
<th>Social Responsibility Aspect</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contributing to the classroom and school community</td>
<td>Being friendly, helpful, cooperative, following classroom rules, and contributing to discussions and activities</td>
</tr>
<tr>
<td>Solving problems in peaceful ways</td>
<td>Resolving conflict by listening to others and drawing upon simple, effective, and appropriate strategies</td>
</tr>
<tr>
<td>Valuing diversity and defending human rights</td>
<td>Demonstrating interest in justice, and treating others fairly and respectfully</td>
</tr>
<tr>
<td>Exercising democratic rights and responsibilities</td>
<td>Showing a growing sense of interest and responsibility in finding ways to make the world a better place</td>
</tr>
</tbody>
</table>

Source. BC Ministry of Education, 2001

There are two rating scales available: an Elaborated Social Responsibility Scale (ESRS) and a Social Responsibility Quick Scale (SRQS). The ESRS provides a one-page document for each aspect of social responsibility by providing detailed descriptions of observable behaviours (i.e., one aspect per page). In contrast, the SRQS provides a one-page document for all four aspects of social responsibility by providing summary descriptions of observable behaviours. These rating scales were created for voluntary use by BC educators so that they may have a
means to assess student social responsibility and use the data collected to inform their instruction (BC Ministry of Education, 2001).

1.3 Universal Strength-Based Screening

In recent years, assessment in educational practice has evidenced a few notable changes. The first notable change is the movement away from traditional service delivery models toward multi-tiered prevention and intervention approaches [e.g., Individuals with Disabilities Education Improvement Act (IDEIA; 2004); McIntosh, MacKay et al., 2011]. The second movement is a shift away from a deficit-based approach to assessment, practice, and research toward one that includes a focus on strengths and competencies (Jimerson, Sharkey, Nyborg, & Furlong, 2004).

1.3.1 From a Traditional to Multi-Tiered Approach

Traditional service delivery models encourage a dichotomous approach to problem identification (e.g., disability presence or absence) and a “wait-to-fail” approach as a condition for the provision of services, which increases the likelihood of negative student outcomes, such as increased aversion to learning, anti-social behaviour, and school dropout (McIntosh, MacKay et al., 2011). In contrast, a multi-tiered prevention and intervention approach provides a continuum of comprehensive assessment and support services, which rely on data generated from two types of formative assessment: (a) universal screening and (b) progress monitoring (McIntosh, MacKay et al.). Universal screening is a practice in which all students at the classroom, grade, and/or school level are periodically assessed throughout the school year to determine whether they are meeting criteria with regard to targeted academic and/or behaviour outcomes. Indicated instructional support is provided for those who are struggling to or do not meet criteria. Progress monitoring assesses whether the identified students are responsive to the increased instructional support. For students who are unresponsive to the increased support, decisions based on progress
monitoring data are used to inform the selection of strategies to better support their development academically and/or behaviourally. In essence, at-risk students are identified and indicated instructional support is provided without delay, and their progress is monitored to inform decisions about withdrawing or providing more specialized support (Hawkins, Barnett, Morrison, & Musti-Rao, 2010).

1.3.2 From a Deficit- to a Strength-Based Approach

Identification of at-risk students can proceed by focusing on deficits (i.e., minimizing/eliminating unwanted outcomes) or on strengths and competencies (i.e., maximizing/promoting desired outcomes). Although the deficit-based approach has proven useful in the diagnosis and treatment of psychological problems, its negative lens has limited investigation into factors that lead to their prevention (Terjesen, Jacofsky, Froh, & DiGiuseppe, 2004). In contrast, by appropriating a positive lens, the strength-based approach expands the conceptualization of optimal human development from a mere absence of symptoms to include the promotion of preventive and protective factors. The strength-based approach places greater emphasis on assets at the individual, family, and community level, rather than deficiencies, to facilitate optimal human development and functioning (Terjesen et al.).

Several advantages of a strength-based approach include: (a) providing a positive context from which to work with children and their families; (b) empowering them to take responsibility for and feel a sense of agency over the outcomes in their lives; (c) enhancing the process of school-based consultation and collaboration; (d) identifying key variables that have large positive effects (“keystone variables”) for use in intervention planning; and (e) promoting more hope and optimism in professionals who work with at-risk and high-needs children and youth and their families (Jimerson et al., 2004). Merrell, Cohn, and Tom (2011) also proposed that
there is an intuitive link between identifying strengths and competencies and targeting them for instructional support in intervention planning, and that this practice facilitates the positive statement of targets for intervention and reduces the stigma associated with assessment.

In sum, integrating strength-based assessment measures into the process of universal screening facilitates the intuitive link between the assessment being conducted and recommended interventions (Merrell et al., 2011). To this end, the abbreviated version of the student social responsibility rating scale created by the BC Ministry of Education, the SRQS, appears to be a measure worth investigating.

1.4 Measuring Social Responsibility with Validity and Reliability

In a multi-tiered model, the primary tier involves universal identification of target variables that should be observable, measurable, and amenable to change (Hawkins et al., 2010). In addition, an important consideration of implementing a universal screening process is whether the data that are being collected are valid and reliable (Severson et al., 2007). Specifically, the screening instrument should adequately measure the construct(s) it purports to measure, and the information it produces should be consistent and stable [The Standards for Educational and Psychological Testing developed jointly by the American Educational Research Association (AERA); American Psychological Association (APA); and National Council Measurement in Education (NCME), 1999]. Furthermore, the screening instrument should be efficient in terms of cost and effort, and solve a high priority problem (e.g., student social responsibility) that is consistent with the school’s core educational objectives to maximize fidelity of implementation and the acceptability of the process to school personnel (Severson et al.). These considerations are fundamental to the success of a multi-tiered approach, as they allow school personnel to answer important questions (e.g., Is the current plan achieving intended outcomes? What areas
need improvement? Which students need additional support?”) and establish a clear and unambiguous link between screening outcomes and recommended instructional and intervention support services (Hawkins et al., 2010; McIntosh, Reinke, & Herman, 2010).

1.5 **Strength-Based Screening Instruments**

Consistent with Severson and colleagues (2007), a key element in maximizing the acceptability of the universal screening process is that it be as efficient as possible. However, Buckley and Epstein (2004) noted that there are few strictly strength-based behaviour assessment tools; many of those that are available embody a deficit- rather than strength-based paradigm.

Three recent reviews of social-emotional assessment tools for children and youth provide guidance for the field. Two reviews focused on available strength-based measures (Jimerson et al., 2004; Merrell & Gueldner, 2010), and one included measures with a strength- and deficit-based focus (Haggerty, Elgin, & Woolley, 2010). Of the measures reviewed, only five were entirely strength-based: (a) Behavioral and Emotional Rating Scale, Second Edition (BERS; Epstein, 2004); (b) Developmental Assets Profile (DAP; Search Institute, 2004); (c) Devereux Student Strengths Assessment (DESSA; LeBuffe, Shapiro, & Naglieri, 2009); (d) Social-Emotional Assets and Resilience Scales (SEARS; Merrell, 2008); and (e) Emotional Quotient Inventory, Youth Version (EQI-YV; BarOn & Parker, 2000). Of these, only one offered an abbreviated version that teachers could use to screen their students in a single classroom period, the DESSA-mini. The DESSA-mini has eight items, whereas the DESSA has 72. These results suggest that although there have been great advances in the development and validation of strength-based assessment tools, more work needs to be done (Merrell et al., 2011). The SRQS created by the BC Ministry of Education represents a potentially valuable strength-based screening instrument; however, although available for voluntary use for BC educators since
2001, its validity and reliability for assessing student social responsibility has not been assessed to date. This study presents the first validation data for the SRQS.

1.6 Purpose of the Study

The purpose of the present study was to evaluate the content validity of a measure of student social responsibility available for voluntary use by BC educators, the SRQS (BC Ministry of Education, 2001). The following research questions were addressed:

1. Content Validity: According to an expert panel, is the content of the SRQS perceived as valid for measuring student social responsibility?

2. Factor Structure: What is the factor structure of the SRQS?

3. Reliability: What is the reliability of the SRQS?
   a. What is the internal consistency reliability of the SRQS ratings?
   b. What is the inter-item and corrected item-total correlation of the items on the SRQS?
   c. What is the test-retest reliability of the SRQS ratings?

4. Construct Validity:
   a. To what extent are teacher ratings of student social responsibility correlated with indices of student problem behaviour?
      i. What is the strength of association between teacher ratings of student social responsibility and individual counts of office discipline referrals (ODRs)?
      ii. What is the strength of association between teacher ratings of student social responsibility and individual counts of suspensions?
b. To what extent do teacher ratings of student social responsibility vary by student characteristics?
   
i. To what extent do teacher ratings of student social responsibility vary by student gender?
   
ii. To what extent do teacher ratings of student social responsibility vary by student grade level?
Chapter 2: Method

2.1 Participants and Settings

Seven individuals identified as subject-matter experts (herein “experts”) evaluated the content validity of the SRQS and its items. The level of expertise and the diversity of knowledge sought for the present study was based on one or more of the following criteria: (a) extensive experience with students; (b) experience in a school-based setting at the district and school level; and/or (c) experience conducting and publishing research on behaviours associated with student social responsibility. Having seven experts in the present study is consistent with the range recommended by the research literature (i.e., between 2 and 20; Gable & Wolf, 1993; Lynn, 1986; Sheatsley, 1983; Walz, Strickland, & Lenz, 1991).

Teachers from two elementary schools (Kindergarten through Grade 6) in a rural school district in British Columbia, Canada completed the SRQS for 360 students (51%). Both schools had been implementing SWPBS for at least 5 years, with one of the schools in its first year of implementing a play-based social-emotional learning (SEL) program. Data collected using the School-wide Evaluation Tool (SET; Sugai, Lewis-Palmer, Todd, & Horner, 2001), a research-validated external evaluation of implementation of the critical features of school-wide behaviour support, indicates how well each school was implementing SWPBS. Research evaluating the SET indicates that schools meeting two key criteria (80% or more for Expectations Taught and 80% or more for Implementation Average) are implementing SWPBS adequately. SET data was not available for the 2011-2012 school year; however, they were for the 2010-2011 school year. The SWPBS-only school met criteria for Expectations Taught (90%), but not for Implementation Average (77%). Similarly, the SWPBS school in its first year of implementing a play-based SEL program met criteria for Expectations Taught (80%), but not for Implementation Average (78%).
Enrollment in the school district’s 20 elementary schools for the 2010-2011 school year was 7,006. Six percent of students in the district were identified as English Language Learners. Consent was obtained from the experts, teachers, and students’ primary caregivers before the start of the present study.

2.2 Measures

2.2.1 Social Responsibility Quick Scale (SRQS)

The SRQS, created by the BC Ministry of Education (2001), captures aspects that broadly reflect accepted expectations and standards of student social responsibility (see Appendix A). The four aspects of social responsibility are: (a) contributing to the classroom and school community; (b) solving problems in peaceful ways; (c) valuing diversity and defending human rights; and (d) exercising democratic rights and responsibilities. The four performance levels are: (a) not yet within expectations; (b) meets expectations (minimum level); (c) fully meets expectations; and (d) exceeds expectations. The rating scale includes descriptive anchors for each performance level to aid raters in accurate and reliable ratings. These anchors vary based on age ranges.

There are four SRQS forms based on different age groups, two of which overlap: (a) Grades K to 3, (b) 4 to 5, (c) 6 to 8, and (d) 8 to 10. The Grades 6 to 8 form is recommended for Grade 8 students in middle school and the Grades 8 to 10 form for Grade 8 students in high school. Age-appropriate examples are provided to facilitate evaluations for each of the forms. Teachers are advised to base their ratings on a number of observations from an adequate sample of different contexts and settings. Ratings are based on teacher observations and the creation of “samples,” which include tasks created by teachers to assess students’ level of social responsibility. A sample describes the performance level attained by a student based on teacher observations and a portion of his or her entire work (e.g., problem solving process, reflections).
To measure the level of student social responsibility, the performance level for each social responsibility aspect was converted to a four-point scale (from 1 = not yet within expectations to 4 = exceeds expectations). Ratings for each aspect were then summed to create a composite score for each student.

2.2.2 Content Validity Survey

A content validity survey was created to obtain information on the SRQS and its items (see Appendix B). The survey included questions regarding the extent to which the SRQS assessed student social responsibility and was developmentally appropriate. Four items assessed to what extent respondents rated each of the four aspects used to define student social responsibility as representative of the construct. There were also six additional questions, two about performance-level classifications (one of which was a four-part question), one about developmental appropriateness, two about clarity of instructions and ease of use of the response format, and one about the SRQS as a whole. The response format for each item was a 4-point Likert-type scale (from 1 = strongly disagree to 4 = strongly agree) and an open-ended field for additional comments.

2.2.3 Office Discipline Referrals (ODRs)

Office discipline referrals (ODRs) are a measure used to capture instances of disruptive problem behaviour in schools (Sugai, Sprague, Horner, & Walker, 2000). Each ODR represents a major incident in which a student violated a behaviour rule, a school staff member identified and reported the violation, and an administrative staff member filled out a report documenting the whole incident. McIntosh, Campbell, Carter, and Zumbo (2009) examined the concurrent validity of ODRs and found statistically significant correlations between ODRs and out-of-school suspensions ($r = .76$), and ODRs and the Externalizing Composite score ($r = .51$) of the
Behavior Assessment Scale for Children–Second Edition (BASC-2; Reynolds & Kamphaus, 2004). To measure the level of student problem behaviour, the number of ODRs received by each participant for the 2011-2012 school year was summed. These data were available from an extant database.

2.2.4 Suspensions

School administrators typically issue in- or out-of-school suspensions when students engage in severe problem behaviour that compromise the ability of school personnel to perform their duties and for other students to learn safely and effectively. Examples of problem behaviour that may result in a suspension include fighting, defying authority, and vandalizing school property (BC Ministry of Education, 1999). McIntosh and colleagues (2009) reported statistically significant correlations between out-of-school suspensions and the Externalizing Composite score \(r = .51\) of the BASC-2 (Reynolds & Kamphaus, 2004). To measure the level of student problem behaviour, the number suspensions (a sum of both in- and out-of-school suspensions) received by each participant for the 2011-2012 school year was summed. These data were available from an extant database.

2.3 Procedures

2.3.1 Content Validity

Ten experts were each sent an email that included a link to an online content validity survey along with electronic copies of the SRQS forms. Of the 10 experts who were invited to participate in the present study, seven completed the survey answering all items. The response rate of 70% is well above the average reported response rate for email surveys (24%; Sheehan, 2001).
2.3.2 Recruitment of Sites

Originally, the two schools in the present study were implementing SWPBS and also involved in a study of the effectiveness of a play-based SEL program. Of the two schools, one had contracted with program developers (not affiliated with the present research team) to receive training for the 2011-2012 school year and implement the program school-wide, and the other, recruited from the same school district using data from the BC Ministry of Education Public School Reports (2011), agreed to act as a comparison and maintain existing practices. Due to a teacher job action in the district, the program was not implemented as intended in the treatment school, but both schools followed through with data collection.

2.3.3 Data Collection

Participating teachers completed the SRQS for each of their students in September 2011 and June 2012. In September 2011, SRQS forms were completed for 54% of students. Due to student mobility, the SRQS was completed for 51% of students in June 2012. A $t$-test was conducted to assess the extent to which students who moved before June differed on SRQS scores, and there were no significant differences ($t = .190, p = .849$). A designated contact at each school distributed and collected the SRQS forms and entered the responses into a spreadsheet without identifying information. The contacts also provided anonymous extant student data (individual counts of ODRs and suspensions) for the 2011-2012 school year.

2.4 Design and Analyses

Data were analyzed using separate procedures for each research question. Only students with SRQS ratings in September 2011 and June 2012 were included in the analyses ($N = 360$). Preliminary analyses were conducted to verify that the data set did not violate the basic assumptions associated with each procedure. SPSS version 20 was used to analyze the data.
2.4.1 Content Validity

A panel of experts completed the content validity survey, providing their feedback on each item of the SRQS and the measure as a whole. Procedures recommended by Rubio and colleagues (2003) were used. The inter-rater agreement (IRA) was calculated to determine the level of agreement between the experts for each item and the SRQS. The IRA was used to calculate the reliability of the experts’ ratings and not the measure itself. The content validity index (CVI), which measures perceptions of the measure’s assessment of the construct, was calculated to quantify these results for each item and the SRQS.

Inter-rater agreement (IRA). As a preliminary analysis, the IRA among the experts was calculated for ratings of each item and the overall rating scale. The rating scale items were dichotomized into two categories: *agree* or *disagree*. All items with a rating of 1 (*strongly disagree*) or 2 (*disagree*) were combined, and those with a rating of 3 (*agree*) or 4 (*strongly agree*) were combined. This method is consistent with recommended practice in the research literature (e.g., Davis, 1992; Grant & Davis, 1997; Lynn, 1986; Rubio et al., 2003). The number of items identified as *agree* were counted, and similarly for those identified as *disagree*. The IRA for each item was determined by calculating the average agreement among the experts. The number of items with an IRA of at least .80 was then divided by the total number of items to calculate the IRA for the overall rating scale, as recommended by Lynn (1986) when the number of experts exceeds five.

Content validity index (CVI). The content validity index (CVI) was calculated for each item and the overall rating scale. The CVI for each item was determined by counting the number of experts who rated an item as *agree* or *strongly agree* and dividing it by the total number of experts. The CVI for the overall rating scale was determined by counting the number of items
rated by the experts as agree or strongly agree and dividing it by the total number of items. A CVI of at least .80 is recommended for new measures (Davis, 1992).

2.4.2 Factor Structure

Using exploratory factor analysis (EFA), following Field’s (2009) recommendations, a principal component analysis (PCA) was used, because “it is concerned only with establishing which linear components exist within the data and how a particular variable might contribute to that component” (p. 638).

Prior to performing the PCA, assumptions underlying factor analysis were tested. To test the suitability of the data, correlation matrices were generated and examined. Coefficient values between .30 and .90 indicate suitability of the data for analysis (Field, 2009). Kaiser-Meyer-Olkin (KMO) values were calculated to measure sampling adequacy. Values between .80 and .90 are considered great (Hutcheson & Sofroniou, 1999). Bartlett’s test of sphericity was calculated to measure whether the overall correlations between variables were significantly different from zero. A significant result indicates that the variables are measuring similar things (Field). Scree plots and eigenvalues were examined to determine factor selection. Common criteria for evaluating factor structure include retaining factors to the left of the point of inflexion on the scree plots and those with eigenvalues greater than 1. The Kaiser criterion is accurate when there are fewer than 30 variables and communalities after extraction (the proportion of shared variance associated with each question) are greater than .70 (Field). A factor loading of at least .40 is typically required of an item for it to be considered to load on a factor (Field).
2.4.3 Reliability

Internal consistency reliability, inter-item and corrected item-total correlations, and test-retest reliability were conducted to assess the stability and consistency of ratings generated by the SRQS.

Internal consistency reliability. The internal consistency reliability of the SRQS was calculated based on teacher ratings of student social responsibility in September 2011 and June 2012. The measure of internal consistency reliability used was coefficient alpha, for which the recommended criterion is .70 for experimental measures and .80 for established measures (Nunnally & Bernstein, 1994).

Inter-item and corrected item-total correlations. The inter-item and corrected item-total correlations were calculated based on teacher ratings of student social responsibility in September 2011 and June 2012. Correlation coefficients between .50 and 1.00 indicate a strong relationship, and the minimum recommended corrected item-total correlation coefficient is .30 (Field, 2009).

Test-retest reliability. Given that the assumptions underlying a paired samples t-test were violated (normality, homogeneity of variance), the Wilcoxon Signed Rank Test, a non-parametric alternative, was used to compare teacher ratings of student social responsibility between the September 2011 and June 2012 SRQS administrations.

2.4.4 Construct Validity

Correlation coefficients were calculated to determine the strength of association between teacher ratings of student social responsibility in June 2012 and indices of student problem behaviour for the entire 2011-2012 school year.
Prior to calculating the correlation coefficients, assumptions underlying Pearson’s correlation were tested. To test for normality (the assumption that scores for each variable are independently and normally distributed), the Kolmogorov-Smirnov statistic was calculated, and the shape of the distribution of scores was examined by inspection of the histograms and the normal Q-Q plots for each variable. To test for homogeneity of variance (the assumption that each variable has equal variability at every level of all other variables), Levene’s test of homogeneity of variance was calculated. To test for linearity (the assumption that the relationship between two variables is linear), a plot of the observed versus predicted values was examined. To test for homoscedasticity (the assumption that the residuals at each level of the predictor variable is constant), a plot of the regression standardized predicted value versus the regression standardized residual was examined. The results indicated violation of all of the tested assumptions. As a result, a non-parametric alternative to Pearson’s correlation was used to calculate the correlation coefficients, Spearman’s Rank Order Correlation (rho). Values ranging from .10 to .29 are considered small, .30 to .49 medium, and .50 to 1.00 large (Field, 2009).

**Teacher SRQS ratings and ODR data.** Teacher ratings of student social responsibility in June 2012 were correlated with total ODRs for the 2011-2012 school year.

**Teacher SRQS ratings and suspension data.** Teacher ratings of student social responsibility in June 2012 were correlated with total suspensions for the 2011-2012 school year.

The extent to which teacher ratings of student social responsibility varies by student characteristics was calculated using non-parametric alternative tests, the Mann-Whitney U Test and Kruskal-Wallis Test, because the assumptions underlying an independent samples t-test and one-way between-groups analysis of variance (ANOVA) were violated (normality, homogeneity of variance).
**Teacher SRQS ratings and student gender.** Median teacher ratings of student social responsibility in September 2011 and June 2012 for male and female students were compared to determine whether they differed significantly in terms of the SRQS ratings they received.

**Teacher SRQS ratings and student grade level.** Median teacher ratings of student social responsibility in September 2011 and June 2012 for students in Grades K to 6 were compared to determine whether they differed significantly in terms of the SRQS ratings they received.
Chapter 3: Results

3.1 Content Validity

3.1.1 Inter-Rater Agreement (IRA)

The agreement of the expert ratings was above the .80 criterion for 12 of the 13 items on the content validity survey (see Table 3.1). The overall IRA was .923. These results indicate that experts provided highly similar ratings regarding the relevance of each item for measuring social responsibility and that the CVI results could be interpreted with confidence.

3.1.2 Content Validity Index (CVI)

The proportion of experts who endorsed each item for measuring social responsibility as content valid ranged from 0 to 100% (see Table 3.1). The CVI for the SRQS as a whole was .923. These results show that 12 of the 13 items on the content validity survey were rated as meeting the .80 criterion. With regard to the item that did not meet criterion, responses from the experts indicated that the developmental appropriateness of the operational definitions for the performance-level classifications for students in Kindergarten to Grade 3 could be improved. One recommendation cited the importance of including examples representative of basic developmental milestones or skills expected of younger students (e.g., turn taking in play, verbal interactions).
Table 3.1  Items from the Content Validity Survey as Rated by Experts

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>IRA</th>
<th>CVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 The four aspects of SR accurately reflect the overall construct of SR</td>
<td>3.714</td>
<td>.488</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2 Each of the four aspects of SR should be included in the instrument</td>
<td>3.429</td>
<td>.535</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3 The instrument should drop one or more aspects of SR</td>
<td>1.714</td>
<td>.488</td>
<td>1</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4 The instrument should add one or more aspects of SR not captured by the instrument</td>
<td>2.286</td>
<td>.756</td>
<td>2</td>
<td>4</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>5 In general, performance-level classifications are valid</td>
<td>3.143</td>
<td>.378</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6 The operational definitions of the performance-level classifications are specific, clear, observable, and appropriate for:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) Kindergarten to Grade 3</td>
<td>2.429</td>
<td>.787</td>
<td>1</td>
<td>3</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>(b) Grade 4 to 5</td>
<td>3.000</td>
<td>.000</td>
<td>3</td>
<td>3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(c) Grade 6 to 8</td>
<td>3.143</td>
<td>.378</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(d) Grade 8 to 10</td>
<td>3.143</td>
<td>.378</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7 Across the grade levels, the instrument is a developmentally appropriate measure of SR</td>
<td>2.857</td>
<td>.378</td>
<td>2</td>
<td>3</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>8 The instructions for using the instrument are clear</td>
<td>3.286</td>
<td>.488</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>9 The response format is easy to use</td>
<td>3.000</td>
<td>.577</td>
<td>2</td>
<td>4</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>10 Overall, this is a valid measure of SR</td>
<td>2.857</td>
<td>.378</td>
<td>2</td>
<td>3</td>
<td>86</td>
<td>86</td>
</tr>
</tbody>
</table>

Note. SR: social responsibility

3.2  Factor Structure

3.2.1  Principal Component Analysis (PCA)

Based on teacher ratings of student social responsibility in September 2011 and June 2012, a principal component analysis (PCA) was conducted. Kaiser-Meyer-Olkin (KMO) values for September 2011 and June 2012 were, respectively, .866 and .865, indicating that the PCAs should yield factors that are distinct and reliable. Bartlett’s test of sphericity $\chi^2 (6)$ values for September 2011 and June 2012 were, respectively, 1185.290 and 1198.408, indicating that correlations between the items of the SRQS were sufficiently large for the PCAs. Eigenvalues were obtained for each component in the data. For September 2011 and June 2012, one
component had an eigenvalue over Kaiser’s criterion of 1 and explained, respectively, 83.365% and 83.563% of the variance. An examination of both scree plots showed one clear inflexion that would justify retaining one component (see Figure 3.1). Given the large sample size ($N = 360$), and the convergence of the scree plots and Kaiser’s criterion on one component, one component was retained in the final analysis. No rotation was conducted, because only one component was extracted. Table 3.2 shows the factor loadings. The items clustered on the one component with factor loadings greater than .40, suggesting that the items are measuring the same factor.

Figure 3.1 Scree Plots for September 2011 and June 2011

<table>
<thead>
<tr>
<th>Items</th>
<th>September 2011</th>
<th>June 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution to the classroom and school community</td>
<td>.922</td>
<td>.921</td>
</tr>
<tr>
<td>Solving problems in peaceful ways</td>
<td>.911</td>
<td>.914</td>
</tr>
<tr>
<td>Valuing diversity and defending human rights</td>
<td>.911</td>
<td>.913</td>
</tr>
<tr>
<td>Exercising democratic rights and responsibilities</td>
<td>.908</td>
<td>.908</td>
</tr>
</tbody>
</table>

### 3.3 Reliability

#### 3.3.1 Internal Consistency Reliability

Based on teacher ratings of student social responsibility in September 2011 and June 2012, alpha coefficients were, respectively, .933 and .934, indicating strong internal consistency reliability
(see Table 3.3). Alpha coefficients if item deleted in September 2011 and June 2012 ranged, respectively, from .908 to .915 and from .910 to .917. These results indicate that deletion of any of the four items of the SRQS would not improve its reliability.

### 3.3.2 Inter-Item and Corrected Item-Total Correlation

Based on teacher ratings of student social responsibility in September 2011 and June 2012, inter-item correlation coefficients ranged, respectively, from .764 to .790 and from .761 to .790. The corrected item-total correlation coefficients in September 2011 and June 2012 ranged, respectively, from .835 to .857 and from .835 to .857. These results indicate that the items on the SRQS are measuring the same factor (see Table 3.3).

**Table 3.3 Inter-Item and Corrected Item-Total Correlations for the SRQS: Coefficient alpha**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Corrected Item-Total Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>September 2011</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>.764</td>
<td>.764</td>
<td>.790</td>
<td>.835</td>
</tr>
<tr>
<td>2</td>
<td>.764</td>
<td>1.00</td>
<td>.777</td>
<td>.787</td>
<td>.840</td>
</tr>
<tr>
<td>3</td>
<td>.764</td>
<td>.777</td>
<td>1.00</td>
<td>.787</td>
<td>.840</td>
</tr>
<tr>
<td>4</td>
<td>.790</td>
<td>.787</td>
<td>.787</td>
<td>1.00</td>
<td>.857</td>
</tr>
<tr>
<td><strong>June 2012</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>.787</td>
<td>.761</td>
<td>.790</td>
<td>.842</td>
</tr>
<tr>
<td>2</td>
<td>.787</td>
<td>1.00</td>
<td>.770</td>
<td>.785</td>
<td>.845</td>
</tr>
<tr>
<td>3</td>
<td>.761</td>
<td>.770</td>
<td>1.00</td>
<td>.793</td>
<td>.835</td>
</tr>
<tr>
<td>4</td>
<td>.790</td>
<td>.785</td>
<td>.793</td>
<td>1.00</td>
<td>.857</td>
</tr>
</tbody>
</table>

### 3.3.3 Test-Retest Reliability

Based on teacher ratings of student social responsibility in September 2011 and June 2012, the median score on the SRQS in September 2011 was 2.750 and in June 2012, 3.000. These results show a statistically significant increase in teacher ratings of student social responsibility from the start to the end of the school year, \( z = 9.319, p < .001 \), with a large effect size (\( r = .49 \)).
3.4 Construct Validity

3.4.1 Teacher SRQS Ratings and ODR Data

Based on teacher ratings of student social responsibility in June 2012 and total ODR data for the 2011-2012 school year, there was a statistically significant correlation, $\rho = -.384$, $n = 360$, $p < .001$.

3.4.2 Teacher SRQS Ratings and Suspension Data

Based on teacher ratings of student social responsibility in June 2012 and total suspension data for the 2011-2012 school year, there was a statistically significant correlation, $\rho = -.259$, $n = 360$, $p < .001$.

3.4.3 Teacher SRQS Ratings and Student Gender

Based on teacher ratings of student social responsibility in September 2011 for male and female students, a Mann-Whitney U test revealed that female students ($Md = 3.00$, $n = 183$) received statistically significant higher SRQS ratings than male students ($Md = 2.25$, $n = 177$), $U = 10410.000$, $z = 5.963$, $p < .001$, with a moderate difference ($\rho = .314$).

Based on teacher ratings of student social responsibility in June 2012 for male and female students, a Mann-Whitney U test revealed that female students ($Md = 3.00$, $n = 183$) received statistically significant higher SRQS ratings than male students ($Md = 2.75$, $n = 177$), $U = 9746.500$, $z = 6.687$, $p < .001$, with a moderate difference ($\rho = .352$).

3.4.4 Teacher SRQS Ratings and Student Grade Level

Based on teacher ratings of student social responsibility in September 2011 for students in each of Grades K to 6, a Kruskal-Wallis Test revealed non-statistically significant differences in SRQS ratings across the seven different grade levels, $X^2 (6, n = 360) = 4.714$, $p = .581$ (see Figure 3.2). Students in Kindergarten and Grade 3 and 4 recorded the lowest median score.
(Md = 2.500). Students in the other four grades recorded median values that ranged from 2.750 to 3.000.

Based on teacher ratings of student social responsibility in June 2012 for students in each of Grades K to 6, a Kruskal-Wallis Test revealed non-statistically significant differences in SRQS ratings across the seven different grade levels, $X^2 (6, n = 360) = 7.248$, $p = .299$ (see Figure 3.2). The Grade 4 students recorded the lowest median score (Md = 2.750). Students in the other six grades recorded the same median value (Md = 3.000).

Figure 3.2 Teacher SRQS Ratings and Grade Level
Chapter 4: Discussion

The purpose of the present study was to assess the validity and reliability of the SRQS, a brief measure of behaviours associated with student social responsibility. Experts rated the SRQS as possessing adequate content validity. Factor analytic and reliability procedures revealed, respectively, that the SRQS possesses a one-factor structure and ratings generated by the SRQS are consistent and stable. Comparison to other indices of student behaviour indicated that the SRQS measures a construct that is inversely related to problem behaviour. Significant group differences in SRQS ratings were found with regard to student gender, but not grade level.

4.1 Content Validity

Results from the content validity survey provide strong overall support for the SRQS as a measure of student social responsibility. Experts provided highly similar ratings with regard to the four aspects’ relevance to the construct of social responsibility. Of note, the experts’ responses indicated that neither adding nor dropping one or more aspects of social responsibility would improve the content validity of the rating scale. Experts also provided highly similar ratings with regard to the clarity of the instructions and ease of use of the response format on the SRQS. Agreement with regard to the performance-level classifications and their developmental appropriateness was consistent in general, with only one area of disagreement, those for students in Kindergarten to Grade 3. A key recommendation was the provision of examples that capture basic developmental milestones or skills expected of younger students. Merrell and Gimpel (1998) remarked that the developmental changes in student skills and foci for instruction, as shown in theory and research, are often overlooked in the development of social skills assessment tools, which often do not use different items based on student developmental level.
Merrell and Gimpel (1998) identified three specific stages of development with regard to the acquisition and development of social skills: (i) early childhood, (ii) middle childhood, and (iii) adolescence. In their definition, early and middle childhood spans age 3 to 6 or 7. In early childhood, important factors to consider in setting expectations and delineating observable behaviours associated with social responsibility include development of the ability to understand that actions are reversible, that perspective-taking expands from an egocentric to include an other object focus, and that “right” versus “wrong” actions extend beyond consideration of immediate consequences. Additional considerations in middle childhood (ages 7-13) include the development of abstract thought, more nuanced cognitive and affective perspective-taking, and following rules as a function of how it serves to bring harmonious relations to a system (e.g., the classroom context). Adolescence (ages 13-17) includes developmental considerations with regard to internalization of social expectations into perspective-taking and decision-making processes, and development of higher-level cognitive and affective abilities needed for empathy.

Validity is not limited to establishing the psychometric soundness of a measure, but necessarily includes consideration of the consequences facilitated by its use (Messick, 1989). Specifically, given that the acquisition of social skills in children and youth proceeds through successive stages of development, inappropriate developmental expectations facilitate inappropriate value judgments and social consequences (Hubley & Zumbo, 2011). To evaluate students in Kindergarten through Grade 3 according to the same criteria (e.g., “can clarify problems and generate and evaluate strategies”) may result in a misalignment between expectations and instructional support. Such a misalignment could lead to a cycle of frustration and failure, whereby the student is deprived of the opportunity to successively develop a sense of agency regarding social-emotional competence, and the teacher a sense of efficacy.
4.2 Factor Structure
Following this first step, factor analytic procedures provided strong evidence with regard to the intended use of the SRQS, to measure the construct of student social responsibility. A factor analysis using a PCA was conducted and identified one linear component in the data, which accounted for most of the variance in SRQS ratings in both September 2011 and June 2012. The four aspects of the SRQS clustered on this one component, with factor loadings indicating that they are all measuring the same factor. The experts expressed that the SRQS is representative of student social responsibility, and the results from factor analytic procedures provide further evidence that it captures a single construct.

4.3 Reliability
Next, reliability evidence was established with internal consistency reliability, and inter-item and corrected item-total correlation data, which indicated that ratings generated by the SRQS are consistent and stable. The data also indicated that deletion of any one of the four aspects of social responsibility would not improve the reliability of SRQS ratings. Test-retest median SRQS ratings found a statistically significant increase in teacher ratings of student social responsibility from the beginning to the end of the school year. This result suggests that the rating scale may be sensitive to changes in actual student behaviours and/or teacher perceptions.

4.4 Construct Validity
Evidence of construct validity revealed an inverse relationship between SRQS ratings and two indices of problem behaviour. Significant associations were found between SRQS ratings and both ODRs and suspensions. No significant differences were found across grade level, which may be due in part to the changes in behaviour criteria for each form. Changes in mean SRQS scores in September 2011 and June 2012 demonstrate parallel grade transitions captured in the
different forms. That is, the scores fall in line with the changeover in forms. As a result, for the Kindergarten to Grade 3 and Grade 4 to 5 forms, there is a pattern of increasing scores. In between the different forms, there is a drop in scores, which suggests that the teachers based their observations on the criteria provided in the forms they were supposed to be using.

In terms of findings regarding gender differences, female students were consistently rated as having higher levels of social responsibility that their male peers. This finding is consistent with the tendency for teachers to rate female students higher than their male peers on measures of social-emotional competence (Merrell & Gimpel, 1998). In addition, the finding is consistent with those from a validation of the teacher report measure of the SEARS, the SEARS-T, which found that the largest gender difference in scores was related to the Responsibility component (Merrell et al., 2011). According to Merrell and Gimpel, environmental stimuli and social consequences, and to a less compelling degree for the authors, biological factors, shape the differential developmental trajectories of males and females. Generalizations have emerged from this and enduring stereotypes whether accurate or misguided, typically have the effect of inevitably shaping actual student behaviours and teacher expectations. The SRQS may need to be revised to account for gender-based differential developmental trajectories if it is to measure social responsibility in male and female students accurately. For example, in social situations, male students are more likely to behave aggressively, whereas female students are more likely to behave cooperatively (Merrell & Gimpel). As a result, it would be more congruent to compare male and female students with a comparison group of individuals of the same gender and age range. Creating gender norms would increase the specificity of the SRQS and the likelihood of the selection and use of appropriate prevention and intervention strategies.
In summary, the data suggest the SRQS is a promising measure of behaviours associated with student social responsibility. However, further validation studies are recommended as the robustness of the identified behaviours on the SRQS as representative of social responsibility, the developmental appropriateness of the operational definitions for the performance-level classifications, and the measure’s sensitivity for use as a screener and progress-monitoring tool remain to be established.

4.6 Limitations

Some limitations of this study are worth noting. Originally the two participant schools in the present study were involved in a study examining the effectiveness of a play-based social-emotional learning program. Both schools had been implementing SWPBS for at least five years. The comparison school was asked to maintain existing practices, while the treatment school had committed to a yearlong implementation of the play-based social-emotional learning program. Due to a teacher job action in the district, the program was not implemented as intended in the treatment school, but both schools followed through with data collection. Participating teachers from both schools completed the SRQS for each of their students. As a result, the first limitation is the inability to determine the presence, absence, and extent of the effects of the play-based social-emotional learning program on student behaviour. A related limitation is the inability to determine if teacher perceptions, and thereby their ratings of student social responsibility, were influenced by the mere presence of the play-based social-emotional learning program and to what degree.

A second limitation is the teacher job action in the district, which narrowed teachers’ responsibilities to the essential service of teaching. As a result, non-essential practices such as the issuance and recording of ODRs and suspensions were likely lower than in a typical school
year. Anecdotal evidence indicates that many discipline practices were exercised by the school principal and/or vice principal, and that s/he would do so on the spot and not through the proper documentation process and procedure stipulated by best SWPBS practice. A third limitation is that the operational definitions for the performance-level classifications may not adequately account for age- and gender-based developmental differences. A fourth limitation is the time interval between completion of the first and second set of SRQS forms. The recommended time interval for test-retest reliability procedures is 1 to 2 weeks (Pedhazur & Schmelkin, 1991). Intervals longer than this interval, as was the case in the present study, are likely to decrease the accuracy of the reliability of scores generated from a measure (Anastasi & Urbina, 1997). A fifth limitation is that the SRQS ratings were not correlated with other comparable strength-based measures and/or their composites that were designed to assess behaviours associated with student social responsibility. A sixth limitation is that the present study did not establish evidence that the SRQS could be used as a screener and progress-monitoring tool. That is, it is unknown if the SRQS would be sensitive enough to detect changes in behaviours associated with social responsibility. A seventh limitation is that the present study did not establish whether higher SRQS ratings were associated with higher academic achievement as measured by grade-point average or standardized test scores. Lastly, the SRQS was created for voluntary use by BC educators and as such only provides information from one informant: the teacher or related educational professionals. Information from one source, although valuable, is likely insufficient to get a comprehensive understanding of a student’s behaviour (Merrell, 2008).

4.7 Implications for Practice and Future Research

Overall, the findings of the present study suggest that the SRQS is a promising measure for gathering teacher-based data on student social responsibility. The identified behaviours on the
SRQS are observable and measurable, and the evidence indicates that it measures behaviours associated with student social responsibility in a manner that is consistent and stable. To this end, it could prove useful for teachers to obtain an approximate sense of whether school-based social-emotional instructional practices are achieving effects in the intended direction (e.g., What is the influence of core SWPBS practice on student social responsibility?). Other noted strengths of the SRQS are that it addresses a high-priority problem in schools (i.e., student behaviour), approaches behaviour assessment from a strength-based perspective, and is efficient to use in terms of time and cost.

Further validation studies, however, should continue to examine the validity and reliability of the SRQS in school-based settings as the robustness of the identified behaviours on the SRQS as representative of social responsibility and the age- and gender-based developmental appropriateness of the operational definitions for the performance-level classifications remain to be established. Conducting these studies while controlling for the effects of other practices and correlating SRQS ratings with other comparable strength-based measures and/or their composites that were designed to assess behaviour associated with student social responsibility (e.g., DESSA-mini, Responsibility component of the SEARS-T) are recommended. In addition, the sensitivity of the measure should also be examined to determine whether it could be used as a screener and progress-monitoring tool, and whether SRQS ratings could be used to drive data-based decision-making processes. It would also be worthwhile to examine the association of SRQS ratings with indices of academic achievement. Moreover, development of screeners for use by other informants (e.g., student, parents) would provide a more comprehensive profile of a student’s behavioural functioning.
References


Appendices

Appendix A: Social Responsibility Quick Scale (SRQS)

Quick Scale: Grades K to 3 Social Responsibility

Directions: Please rate this student on each of the four aspects of social responsibility by selecting only one response option per aspect. Ratings should be based on your experiences with the student from the beginning of the school year.

1. CONTRIBUTING TO THE CLASSROOM AND SCHOOL COMMUNITY
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

2. SOLVING PROBLEMS IN PEACEFUL WAYS
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

3. VALUING DIVERSITY AND DEFENDING HUMAN RIGHTS
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

4. EXERCISING DEMOCRATIC RIGHTS AND RESPONSIBILITIES
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations
Quick Scale: Grades 4 to 5 Social Responsibility

Directions: Please rate this student on each of the four aspects of social responsibility by selecting only one response option per aspect. Ratings should be based on your experiences with the student from the beginning of the school year.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Not Yet Within Expectations</th>
<th>Meets Expectations (Minimal Level)</th>
<th>Fully Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CONTRIBUTING TO THE CLASSROOM AND SCHOOL COMMUNITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. SOLVING PROBLEMS IN PEACEFUL WAYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. VALUING DIVERSITY AND DEFENDING HUMAN RIGHTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. EXERCISING DEMOCRATIC RIGHTS AND RESPONSIBILITIES</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quick Scale, Version 1.0, Time 1: August 22, 2011
Quick Scale: Grades 6 to 8 Social Responsibility

**Directions:** Please rate this student on each of the four aspects of social responsibility by selecting only one response option per aspect. Ratings should be based on your experiences with the student from the beginning of the school year.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Not Yet Within Expectations</th>
<th>Meets Expectations (Minimal Level)</th>
<th>Fully Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTRIBUTING TO THE CLASSROOM AND SCHOOL COMMUNITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOLVING PROBLEMS IN PEACEFUL WAYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VALUING DIVERSITY AND DEFENDING HUMAN RIGHTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXERCISING DEMOCRATIC RIGHTS AND RESPONSIBILITIES</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Aspect Details:**

- **CONTRIBUTING TO THE CLASSROOM AND SCHOOL COMMUNITY**
  - often appears to be unfriendly and negative
  - does not take responsibility or work cooperatively

- **SOLVING PROBLEMS IN PEACEFUL WAYS**
  - unwilling or unable to solve interpersonal problems; may be illogical or blame others, or become violent or sarcastic
  - tends to view problems in black and white; has difficulty considering more than one perspective, generating strategies, and predicting consequences

- **VALUING DIVERSITY AND DEFENDING HUMAN RIGHTS**
  - often disrespectful and may avoid or be negative towards those perceived as different in some way
  - usually respectful to others and accepting of differences, but may not see the need for action on human rights

- **EXERCISING DEMOCRATIC RIGHTS AND RESPONSIBILITIES**
  - tends to be egocentric or apathetic; may show a sense of powerlessness
  - shows some interest in making the world a better place, but ideas tend to be very general and follow through tends to be inconsistent

Quick Scale, Version 1.0, Time 1: August 22, 2011
Quick Scale: Grades 8 to 10 Social Responsibility

Directions: Please rate this student on each of the four aspects of social responsibility by selecting only one response option per aspect. Ratings should be based on your experiences with the student from the beginning of the school year.

1. CONTRIBUTING TO THE CLASSROOM AND SCHOOL COMMUNITY
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

2. SOLVING PROBLEMS IN PEACEFUL WAYS
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

3. VALUING DIVERSITY AND DEFENDING HUMAN RIGHTS
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

4. EXERCISING DEMOCRATIC RIGHTS AND RESPONSIBILITIES
   - Not Yet Within Expectations
   - Meets Expectations (Minimal Level)
   - Fully Meets Expectations
   - Exceeds Expectations

Quick Scale, Version 1.0
Appendix B: Content Validity Survey

AN INVESTIGATION OF SOCIAL RESPONSIBILITY: THE SRQS

This survey is designed to assess the content validity of the Social Responsibility Quick Scale (SRQS), a scale measuring behaviours related to social responsibility, defined as “a wide and varied range of competencies and dispositions,” including contributing to the classroom and school community; solving problems in peaceful ways; valuing diversity and defending human rights; and exercising democratic rights and responsibility (BC Ministry of Education, 2001, p. 4). The scale has been available since 2001 for voluntary use by educators in BC. Your responses will be used to assess the validity of the scale. Please download the attached Social Responsibility Quick Scales for Kindergarten to Grade 3, Grade 4 to 5, Grade 6 to 8, and Grade 8 to 10 and review them before answering each of the following questions.

Thank you for providing your expertise on this important matter.

Please answer the following about the SRQS:

1. The four aspects of social responsibility (contributing to the classroom and school community; solving problems in peaceful ways; valuing diversity and defending human rights; and exercising democratic rights and responsibility) accurately reflect the overall construct of social responsibility.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Additional comments:

2. Each of the four aspects of student social responsibility should be included in the instrument.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Additional comments:

3. The instrument should drop one or more aspects of social responsibility from the instrument.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Please specify:

4. The instrument should add one or more aspects of social responsibility not captured by the instrument.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Please specify:

5. In general, the performance-level classifications (not yet within expectations; meets expectations (minimum level); fully meets expectations; exceeds expectations) are valid.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   Additional comments:
6. Specifically, the operational definitions of the performance-level classifications (i.e., the bullet points providing examples of not yet within expectations; meets expectations (minimum level); fully meets expectations; exceeds expectations) are specific, clear, observable, and appropriate for:

a. Kindergarten to Grade 3
   Strongly Disagree  Disagree  Agree  Strongly Agree

b. Grade 4 to 5
   Strongly Disagree  Disagree  Agree  Strongly Agree

c. Grade 6 to 8
   Strongly Disagree  Disagree  Agree  Strongly Agree

d. Grade 8 to 10
   Strongly Disagree  Disagree  Agree  Strongly Agree

Additional comments:

7. Across the grade levels, the instrument is a developmentally appropriate measure of social responsibility.

a. Strongly Disagree  Disagree  Agree  Strongly Agree

Additional comments:

8. The instructions for using the instrument are clear.
   Strongly Disagree  Disagree  Agree  Strongly Agree

Additional comments:

9. The response format is easy to use.
   Strongly Disagree  Disagree  Agree  Strongly Agree

Additional comments:

10. Overall, this is a valid measure of social responsibility.
    Strongly Disagree  Disagree  Agree  Strongly Agree

Additional comments:
Please answer the following demographic questions:

11. What is your current primary role?

   Classroom teacher (general education)   Graduate student researcher
   Resource/learning assistance teacher    University faculty researcher
   District personnel                     Other

   If “other,” please specify:

12. How many years have you been in this position?

13. How many years have you been working in education?

14. What is your ethnicity? (Options based on the Canada Census form)

   North American Indian
   Métis
   Inuit (Eskimo)
   White
   Chinese
   South Asian (e.g., East Indian, Pakistani, Sri Lankan etc.)
   Black
   Filipino
   Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian etc.)
   Arab
   West Asian (e.g., Iranian, Afghan etc.)
   Korean
   Japanese
   Other

   If “other,” please specify:

15. What is your gender?

   Male   Female

Thank you very much for sharing your time and expertise with us!