School Psychology Quarterly

Effects of Mischievous Responding on Universal Mental Health Screening: I Love Rum Raisin Ice Cream, Really I Do!

Michael J. Furlong, Aileen Fullchange, and Erin Dowdy


CITATION
Effects of Mischievous Responding on Universal Mental Health Screening: I Love Rum Raisin Ice Cream, Really I Do!

Michael J. Furlong, Aileen Fullchange, and Erin Dowdy
University of California, Santa Barbara

Student surveys are often used for school-based mental health screening; hence, it is critical to evaluate the authenticity of information obtained via the self-report format. The objective of this study was to examine the possible effects of mischievous response patterns on school-based screening results. The present study included 1,857 high school students who completed a schoolwide screening for complete mental health. Student responses were reviewed to detect possible mischievous responses and to examine their association with other survey results. Consistent with previous research, mischievous responding was evaluated by items that are legitimate to ask of all students (e.g., How much do you weigh? and How many siblings do you have?). Responses were considered “mischievous” when a student selected multiple extreme, unusual (less than 5% incidence) response options, such as weighing more than 225 pounds and having 10 or more siblings. Only 1.8% of the students responded in extreme ways to 2 or more of 7 mischievous response items. When compared with other students, the mischievous responders were less likely to declare that they answered items honestly, were more likely to finish the survey in less than 10 min, reported lower levels of life satisfaction and school connectedness, and reported higher levels of emotional and behavioral distress. When applying a dual-factor mental health screening framework to the responses, mischievous responders were less likely to be categorized as having complete mental health. Implications for school-based mental health screening are discussed.

Keywords: screening, schools, complete mental health, mischievous response, data screening

Universal screening for complete mental health has been proposed as a key step in service delivery reform, with the potential to place increased emphasis on prevention, early intervention, and promotion of social–emotional health (Dowdy et al., 2014; Moore et al., 2015). Schools are the primary place in which children receive mental health services, and for some, it is the only source of mental health care (Burns et al., 1995; Stiffman et al., 2000). However, multiple studies indicate that children who score high on ratings of clinical symptomology often do not receive any services, and this is particularly true for minority populations (Burns et al., 2004; Kataoka, Zhang, & Wells, 2002; Leaf et al., 1996), perhaps indicating that individuals who qualify for services are not being identified in school settings. Despite the need for mental health services and research showing linkages between social–emotional well-being and academic functioning (Becker, Brandt, Stephan, & Chorpita, 2014; Strolin-Goltzman, Sisselman, Melekis, & Auerbach, 2014), schoolwide mental health screening is a relatively recent phenomenon, with approximately 12% of schools engaged in this practice (Bruhn, Woods-Groves, & Huddle, 2014). However, universal screening is an important component of efforts to increase access to needed services. With the recent reauthorization of the Elementary and Secondary Education Act, which now pro-
vides funding for schools to expand mental health services, including assessing students for eligibility of such services (www2.ed.gov/policy/elsec/leg/esea02/pg/79.html), the issue of mental health screening in schools has been placed in the national spotlight. Universal mental health screening is furthered by popular support: 83% of parents and two thirds of voters support delivery of mental-health-related services in school settings (Chamberlin, 2009).

Data on youth psychosocial functioning are usually collected via questionnaires given to parents, teachers, and/or the youths themselves. Levitt, Saka, Romanelli, and Hoagwood’s (2007) review of instruments used for mental health screening in schools indicated that the majority of instruments included a self-report version or were self-report solely, especially for youths at the secondary grade levels. There must be an assumption, in order to regard data as valid, that youths’ responses are honest. Given the high stakes of incorrectly identifying students for mental health services, it is important that the youths who are identified as high-risk are truly so.

There is evidence that the honesty of survey responses may be at times called into question—that is, social desirability, careless/unengaged responding, and response inconsistencies (Cornell, Klein, Konold, & Huang, 2012). In one study, Furlong, Sharkey, Bates, and Smith (2004) examined responses from the 2001 Youth Risk Behavior Surveillance Survey to see if unusual response patterns were associated with different health-related behavior endorsement rates. They compared students who reported carrying a gun to school six or more times in the previous month (the most extreme possible response) to a matched group of peers with no or low reported school gun possession. The results showed that the extreme gun possession group was more likely to demonstrate implausible response patterns. For example, 5.3% reported consuming alcohol on the school campus “30 or more days” and using marijuana “40 or more times” on a school campus in the previous month. However, this same group was more likely than the comparison students to eat carrots four or more times a week and drink four or more glasses of milk a day, and 33% reported that they engaged in daily aerobic exercise, a highly unlikely combination of behaviors. Furlong et al. (2004), concluded that the frequent school gun possession group, at least in part, was comprised of students who purposely responded with the most extreme response option (good or bad).

To address the issues raised by the Furlong et al. (2004) study, current best practices suggest the use of multiple quality checks in order to ameliorate some of these response authenticity concerns (Cornell, Lovegrove, & Baly, 2014; Meade & Craig, 2012). Additionally, there are a growing number of recent studies on mischievous responders, “youths who provide extreme, and potentially untruthful, responses to multiple questions” (Robinson-Cimpian, 2014, p. 171). Mischievous responding is just one type of response pattern that could affect authenticity of responses. It can be categorized as purposeful faking or content responsive faking, whereas careless responding is characterized as content nonresponsivity, or random responding without regard to the item content (Meade & Craig, 2012). Mischievous or “jokester” responses seem to be associated with responses that indicate high-risk behaviors. For example, Robinson-Cimpian (2014) used screener items to identify extreme responders, such as those reporting being “blind and deaf and in a gang and parenting multiple children” and found that the combination of unlikely responses was more likely to be associated with responses indicating minority status, such as identifying as transgender or disabled. The disparities were quite large, with 40% of supposed transgender-identifying youths providing at least three low-frequency responses versus only 1.5% of cisgender youths. Removal of these mischievous responses and reanalysis of these data eliminated the disparity between trans and cisgender youth in terms of frequency of suicidal thoughts, going against findings from other studies that reported a disparity in suicidal ideation between these two populations (Eliason, 2010). In another study examining data from the National Longitudinal Study of Adolescent Health (Add Health), Fan and colleagues (2002) found that 19% of youths who claimed to be adopted actually had parents who identified their children as biological. Removing the mischievous responses resulted in a reduction or elimination of almost all disparities between adoptees and nonadoptees. In a similar study, it was found that 99% of youths who stated that they had an
artificial limb actually did not, as verified by in-person follow-up interviews; such mischievous responders were also more likely to claim that they were adoptees and have disabilities (Fan et al., 2006). Although Fan and colleagues (2006) chose to follow up with respondents on the basis of the response to one particular item, it is important to note that the literature has not identified any single item as being ideal for identifying mischievous responders. Rather, as in Robinson-Cimpian’s (2014) study, identification of mischievous responding is based on responses from a combination of items that, individually, may be plausible even if unlikely, but, when taken together, become most likely implausible.

**Universal Mental Health Screening**

Although few schools are currently engaged in universal mental health screening, the trend to gather screening data is on the rise (Bruhn et al., 2014). However, current school-based screening approaches often rely heavily on the assessment of mental health problems or risk, without fully considering the positive contributions to mental health (Moore et al., 2015). A recent mental health model that is being applied in, and resonates with, best school-based mental health practices is the dual-factor or dual-continua model, which considers both positive and negative indicators of mental health (e.g., Keyes, 2005). Greenspoon and Saklofske (2001) proposed that a dual-factor assessment model should incorporate both positive (i.e., subjective well-being) and negative (i.e., internalizing and externalizing symptoms) measures. The crossing of these measures are used to create four logical mental health groups: (a) Group 1 (high positive and low negative; complete mental health), (b) Group 2 (low positive and high negative; troubled), (c) Group 3 (low positive and low negative; vulnerable), and (d) Group 4 (high positive and high negative; symptomatic but content). Subsequently, studies using the dual-factor model have assigned all students into one of Greenspoon and Saklofske’s (2001) four dual-factor mental health groups (Antaramian, Huebner, Hills, & Valois, 2010; Kelly, Hills, Huebner, & McQuillin, 2012; Lyons, Huebner, Hills, & Shinkareva, 2012; Suldo & Shaffer, 2008). These studies have examined the incidence of dual-factor categories and how the categories covary with quality of life indicators. Suldo and Shaffer (2008) found group differences, with students in the optimal/complete mental health group having the best reading skills, school attendance, academic self-perceptions, academic-related goals, social support from peers and parents, and self-perceived physical health, with fewer social problems than their peers. Likewise, Antaramian et al. (2010) reported group differences in student engagement, academic achievement, and environmental support for learning. Given the relevance of the dual-factor model to understanding youths’ complete mental health, including an assessment of both strengths and distress and its use for mental health screening (Moore et al., 2015), this present study investigated the influence of mischievous responding on dual-factor classifications derived from student self-reported responses to a universal complete mental health screening survey. This appears to be the first study to investigate the relations between mischievous responders and dual-factor groups.

**Study Purpose**

The aforementioned mischievous responder literature addressed assessments of psychosocial functioning within the context of collecting data for public health assessment. This exploratory study sought to make a contribution by evaluating mischievous responding, a recent perspective that has not yet been applied to school-based universal mental health screening. The following research questions were examined:

1. What proportion of the sample endorsed items designed to detect mischievous responding?
2. How is mischievous responding associated with response honesty and survey relevance?
3. Did mischievous responders answer survey items differently than nonmischievous responders?
4. Using a dual-factor classification model, how did mischievous responding influ-
ence complete mental health screening outcomes?

Method

Participants

The sample comprised 1,857 students (female 51.2%, male 47.7%, other gender self-designation 1.2%) enrolled in a central California comprehensive high school in the following grades: ninth (25.0%), tenth (29.0%), eleventh (25.5%), and twelfth (20.5%). The students' self-identified ethnicity was as follows: Latino/a (45.2%), White (38.5%), multiple group identification (11.0%), Asian (2.9%), Black (1.2%), American Indian (0.5%), Pacific Islander (0.3%), and other (0.4%). The sample represented 86% of the students attending the high school. The participating school had an academic performance index rank of 6 (better than 60% of all state high schools), with 21% of the students designated as English language learners.

Measures

The measures used in this investigation included (a) mischievous response items, (b) items that examined students' survey orientation (attitudes toward taking and valuing the survey), (c) two measures used to create complete mental health groups, and (d) two measures of quality of life (school connectedness and covitality) to examine possible differences between mischievous responders and other students.

Mischievous responding. A primary objective when crafting a mischievous response index is to have a set of items that have low endorsement rates and are relatively neutral and factual. That is, the items are not so far-fetched as to be unbelievable or bogus (e.g., “I can leap tall buildings in a single bound” or “Ironman is a real person”). All items (a) were legitimate questions that could be posed to any student, (b) included plausible and implausible response options, and (c) provided students an opportunity to give exaggerated responses. Research to explore the possible effects of mischievous responding on youth self-report surveys is not yoked to any specific item. It is true that any single item might be true for a given student (e.g., have 10 or more siblings); nonetheless, in the mischievous responding paradigm it is the overall pattern of endorsing low frequency response options that is of interest.

Six items were used in this study, with seven mischievous responses coded. Item responses were considered mischievous if they were unusual, as indicated by less than 5% incidence. Four item indicators (weight, deafness/hearing impairment, blindness, and time since last dental visit) were previously used in the Robinson-Cimpian (2014) study. Two additional factual items with content relevant to the school context (times voted student of the month) and family of origin (number of siblings) were included. Per Robinson-Cimpian’s suggestion, we included an item that could be considered “funny” (ice cream preference). The mischievous response items included and the seven responses options coded as unusual or “mischievous” were as follows:

1-2. Do you have a disability? (check all that apply): hearing impairment, deafness, visual problems, blindness, health impairment, learning disability, none: [1] hearing impairment/deafness and [2] blindness each coded as mischievous (1.6% endorsement).

3. When was the last time you were seen by a dentist? less than one year, one year, two years, three–four years, five or more years: [3] three–four years/five or more years coded as mischievous (3.3% endorsement).

4. How much do you weigh without your shoes on? (sliding scale response format 0–300 lbs.): [4] ≤ 89 (bottom 2.5% of sample) and ≥225 (top 2.5% of sample) coded as mischievous.

5. How many times have you been recognized as a student of the month? 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10+: [5] 10 + coded as mischievous (3.1% endorsement).

6. How many brothers and sisters do you have? 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10+: [6] 6–10+ coded as mischievous (3.2% endorsement).

7. If you had to choose just one flavor of ice cream for one year, which one would you select? chocolate, vanilla, rum raisin, strawberry, I’d go without ice cream: [7] rum raisin coded as mischievous (2.5% endorsement).
All keyed responses (0–7) had less than five percent endorsement (range = 1.2% [blindness] to 4.9% [weight]) and were consistent with Robinson-Cimpian’s (2014) recommendations for mischievous response items.

Survey orientation. We used two items to further explore the general response honesty (self and peers) of students. One has been used in the California Healthy Kids Survey (WestEd, 2014) for more than 15 years. Placed near the end of the survey, students are asked, “How many questions in this survey did YOU answer honestly?” with response options of all, most, some, hardly any, and none. Cornell et al. (2012) found that students acknowledge providing misleading, dishonest responses. We used a second honesty item: “How many OTHER STUDENTS at your school do you think answered the questions in this survey honestly?” The response options were all, most, some, hardly any, or none.

A second brief scale measuring survey relevance was constructed for this study. It consisted of three items that appeared at the end of the survey. The items asked students their opinions about the purpose and potential value of the survey they had just taken. The items were preceded by the following introduction: “Please answer the following questions so that the school staff will have a better understanding about how relevant and important these questions were to your own and other students’ life experiences.” The three items were (a) “These questions helped me to privately tell adults about what is going on in my life,” (b) “The questions in this survey will help my school to better support all students, and (c) “The questions in this survey were relevant to all students’ life experiences.” The response options were strongly disagree, moderately disagree, mildly disagree, mildly agree, moderately agree, and strongly agree. Exploratory factor analysis in Mplus version 7 (Muthén & Muthén, 1998–2011) using principle factor axis extraction was used to evaluate if it was reasonable to combine the three items into one index. One factor was extracted (eigenvalue 2.15) that explained 71.5% of the common variance with items loadings between .69 and .81. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (.71) and the Bartlett’s Test of Sphericity, $\chi^2 = 1735.19$ (3), $p < .001$, supported the unidimensionality of these three items. The alpha reliability coefficient for the present sample was .80.

Quality of life. Two quality of life measures relevant to the school-based survey were employed to examine if mischievous responders produced different mean responses compared with their nonmischievous responding peers, as was previously reported by Robinson-Cimpian (2014).

The five-item School Connectedness Scale is one of the most widely used research measures of student bonding and school attachment (e.g., An- derman, 2002; Libbey, 2004; McNeely, Nonne-maker, & Blum, 2002; Resnick et al., 1997). It uses a five-point response scale: 1 = strongly disagree, 2 = disagree, 3 = neither disagree nor agree, 4 = agree, and 5 = strongly agree, with a total score ranging from 5 to 25. Evidence supports its use as a unidimensional scale with factorial invariance across genders and sociocultural groups (Furlong, O’Brennan, & You, 2011). The alpha reliability coefficient for the present sample was .86.

The Social Emotional Health Survey-Secondary (SEHS-S) is a 36-item instrument that assesses core self–other cognitive dispositions (belief in self, belief in others, emotional competence, and engaged living) that help a student to foster positive development and protect against psychological distress. The positive psychological dispositions’ combined and interactive effects is called covitality (Lenzi, Dougherty, Furlong, Sharkey, & Dowdy, 2015; You et al., 2014). The SEHS-S is validated across multiple, diverse samples, with evidence provided for its higher-order structural invariance model across U.S. (You, Furlong, Felix, & O’Malley, 2015), Korean (Lee, You, & Furlong, 2016), Japanese (Ito, Smith, You, Shimoda, & Furlong, 2015), and Australian (Pennell, Boman, & Mer- gler, 2015) samples. The SEHS-S’ overall total covitality score had an alpha reliability coefficient of .94 for this study’s sample.

Complete mental health. Research applications of the dual-factor model have evaluated youths’ complete mental health by the juxtaposition of scores on a measure of psychological distress and a measure of global life satisfaction. The present study used the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Brief Multidimensional Students’ Life Satisfaction Scale (BMSLSS; Seligson, Huebner, & Valois, 2003).

The SDQ is a widely used brief questionnaire that measures children’s emotional experiences
and behaviors. Through self-report, respondents can select not true, somewhat true, or certainly true for each item. Psychometric analyses of the SDQ have found less than optimal internal consistency and factorial invariance of the original five-factor structure; hence, alternative models have been examined (e.g., Ruchkin, Jones, Vermeiren, & Schwab-Stone, 2008; Stevanovic et al., 2015). In the present study, we used an adapted version of the SDQ that utilizes two of three factors from the three-factor version of the SDQ validated by Ruchkin et al. (2008) with urban and suburban youth in the United States. The third factor, prosocial behavior/peer competence, was not included in this study, as the purpose of the SDQ in this study was to assess the distress/deficit side of youths’ mental health. The SDQ was chosen because of the superior reliability and theoretical relevance of the resulting scales with respect to the purpose of the present study. Consistent with the dual-factor mental health model employed in previous studies and to contribute to the overall measurement efficiency of the survey used in an applied school screening context, especially in regards to survey length, we used the five items with the highest loadings from the confirmatory factor analysis conducted by Ruchkin et al. (2008) for the emotional distress/withdrawal factor (a measure of internalizing problems; original SDQ items 3, 6, 8, 13, and 16; loadings .47–.60; e.g., “I worry a lot”) and the behavioral reactivity/conduct problems factor (a measure of externalizing problems; original SDQ items 2, 5, 10, 15, and 22; loadings .56–.62; e.g., “I get very angry and often lose my temper”). The internal consistency coefficients for emotional distress/withdrawal (.77) and behavioral reactivity/conduct problems (.75) for this study’s sample exceeded the accepted benchmark of .70 (Leary, 2004). To prepare the data to create the dual-factor classifications and to assist with interpretation, the SDQ scores were transformed to T scores using the means and standard deviations reported by Ruchkin et al. (2008), with T scores above 60 indicative of behavioral and/or emotional risk. The norms from the full measure provided by Ruchkin et al. (2008) were considered appropriate given that the study population used by Ruchkin et al. (2008) reflected that of students in the secondary grade levels in the public school system. Our modifications did not change the meaning of items, and there were identical instructions and response categories used in the shortened version of the SDQ as in the original version.

The BMSLSS is a five-item self-report measure that assesses life satisfaction in five domains (one item per domain): school, family, friends, self, and living environment. In this study, a five-point response scale (1 = very dissatisfied to 5 = very satisfied) suggested by Athey, Kelly, and Dew-Reeve (2012) was used and summed across the five items to form a total subjective well-being (SWB) total score (range 5–25). The BMSLSS has been used in research with children from ages 8 to 18 (Huebner, Drane, & Valois, 2000; Huebner, Valois, Paxton, & Drane, 2005). Alpha coefficients for the BMSLSS total SWB index generally fall within the .70–.80 range, and it has been shown to have a unidimensional factor structure (Huebner & Hills, 2013; Huebner, Seligson, Valois, & Suldo, 2006). The alpha reliability coefficient for this study’s sample was .79.

To create mental health categories, we mimicked procedures commonly used within dual-factor approaches to classification (e.g., Suldo & Shaffer, 2008). High and low BMSLSS (1) and SDQ (2) scores were applied as follows to place students into one of the four dual-factor categories proposed by Suldo (2016):

1. **Vulnerable**: (1) BMSLSS score ≤30th percentile of sample and (2) SDQ Emotional/Withdrawal: T score <60 and SDQ Behavior/Conduct T score <60;
2. **Troubled**: (1) BMSLSS score ≤30th percentile of sample and (2) SDQ Emotional/Withdrawal T score ≥60 or SDQ Behavior/Conduct T score ≥60;
3. **Complete Mental Health**: (1) BMSLSS within the top 70th percentile of sample and (2) SDQ Emotional/Withdrawal T score <60 and SDQ Behavior/Conduct T score <60;
4. **Symptomatic but Content**: (1) BMSLSS within the top 70th percentile of sample and (2) SDQ Emotional/Withdrawal T score ≥60 or SDQ Behavior/Conduct T score ≥60.

**Procedures**

Data were collected as part of universal mental health screening carried out at one compre-
hensive high school. Consent forms were sent to the parents or guardians of all enrolled students in the annual enrollment packet distributed prior to the start of school. Following the university’s research review committee approval, passive parental consent and student assent were attained prior to survey administration. During the first month of the 2015–2016 school year, students with parental consent and individual assent completed a universal screening survey via an online Qualtrics format in the school computer lab (56 parents declined consent). Surveys were collected over a 2-week period. For students who were absent from class at the time that their class completed the screening, five additional follow-up attempts were made to allow these students the opportunity to complete the survey. In support of the school’s efforts to identify and support all students’ mental health needs, the students were asked to provide their school identification number prior to completing the survey. Students were asked to provide assent to take the survey and were given the option to privately opt out of the survey (55 students opted out). In addition to the students who declined assent and whose parents declined consent, some students opted out of the survey by not coming to the computer lab. Additionally, some students were not included in the final sample due to chronic truancy or absences on the screening days; having severe disabilities and the school staff noting that the survey was not appropriate for them; and changes in status (e.g., no longer enrolled) from the beginning to the end of data collection. The school attempted to gather screening data from as many students as possible; however, some students were not accounted for despite these efforts.

The measures used in this study were presented in the following order with one item displayed on the monitor at a time within each subscale block: SEHS-S, SDQ, School Connectedness, BMSLSS, and Survey Orientation. Having a consistent survey order for all respondents is similar to the procedure commonly used by schools when administering traditional paper-and-pencil screening surveys. However, in the present study an enhancement over traditionally formatted hard copy surveys was that item presentation was randomized within each individual survey block. Furthermore, the mischievous response items were randomly distributed throughout the middle two thirds of the survey. The online survey was formatted to request, but not require, students to complete an item if they had not responded to (skipped) an item. For the measures used in this study there were no missing data.

Results

Description of Mischievous Item Response Patterns

Most of the students (n = 1,524; 83.5%) endorsed zero mischievous response items (designated MR-0 group). Of the 379 times that a mischievous response item was endorsed, 253 (13.9%) were by students who endorsed just one item (designated MR-1 group). Only 49 (1.8%) students endorsed more than one item (33 endorsed two items, 10 endorsed three items, and 6 endorsed four to six items [designated MR-2 group]). No student endorsed all seven items.

Most students (97.3%) completed the survey in 10 or more minutes, which was reasonable for the 90-item survey (a minimum of 8–10 s per item). A total of 10.2% of the MR-2 + students completed the survey in 5–10 min (no surveys were completed in less than 5 min), as opposed to 3.5% of the MR-0 and 3.2% of the MR-1 students, $\chi^2(2, 1826) = 6.34, p = .042, \varphi = .059$. A greater percentage of students with more mischievous responses completed the survey in 5–10 min than students with less mischievous responses; however, examination of the effect size suggests that the differences are practically negligible.

Mischievous Responding and Survey Orientation

We examined survey orientation by asking students to report on self- and peer-response honesty. As shown in Table 1, students in all groups reported more honesty than dishonesty. However, more MR-0 students reported that they had completed “all” or “most” items honestly compared with the MR-1, and MR-2 + students. Students in all groups perceived less survey taking honesty in their peers, with less than 50% of students expressing the opinion that “all” or “most” of their peers had answered the questions honestly. We also compared re-
responses to the three survey relevance questions. A comparison of means (see Table 2, “Survey relevance” row) showed that the MR-0 and the MR-1 groups did not differ from each other and expressed a more positive opinion about the pertinence and relevance of the survey items than the MR-2 students (\(d = .39\) and .35, respectively, small effect size differences). On average, the MR-0 and MR-1 students’ responses were between the mildly agree to moderately agree response options, reflecting positive relevance toward the survey content. The MR-2 students’ responses were in the mildly disagree to mildly agree response range, indicating a less positive evaluation of the survey purpose.

### Effects of Mischievous Responding on Student Responses

Previous research has reported that mischievous responders provide different responses to nonmischievous questions than their peers. We examined the replication of this finding for the present sample using items that were not used in forming the dual-factor complete mental health groups. As shown in Table 2 (“Quality of life indicators” rows), the MR-2 + students had significantly lower scores than the MR-0 and the MR-1 students on measures of school connectedness and social emotional covitality (all small effect size differences). Also shown in Table 2, a similar pattern was found when the complete mental health dual-factor measures were considered. The MR-0 and the MR-1 group mean scores for BMSLSS, SDQ Emotional/Withdrawal, and SDQ Behavior/Conduct did not differ from each other; in contrast, they both differed significantly from the MR-2 + students (small and medium effect size differences, .24 and .60).

### Mischievous Responders and Dual-Factor Classifications

Given that the MR-2 + group differed from the MR-0 and MR-1 students on the indicators examined, we explored the relations between mischievous responding and the complete mental health status of students by applying Suldo’s (Suldo, 2016; Suldo & Shaffer, 2008) dual-factor mental health classification approach. Applying these criteria resulted in the following dual-factor classifications for the combined sample: 16.2% of the sample was placed in the Vulnerable group, 12.7% in the Troubled group, 62.8% in the Complete Mental Health group, and 8.4% in the Symptomatic but Content group (percentages of 21.1%, 12.1%, 65.0%, and 10.1%, respectively, as reported in previous studies; Suldo, 2016).

The MR-2 + group’s differing means on the BMSLSS and SDQ measures reflect, at least in part, a tendency to exaggerate responses toward the less socially desirable end of a scale’s continuum: lower connectedness, lower life satisfaction, and higher emotional and behavior distress. When considering these responses within a dual-factor framework, this combination would make it more likely that these mischievous responders would exhibit a troubled dual-factor classification pattern. As shown in Table 3, this outcome was found in that 18.4% of the MR-2 + group were in the Troubled classification versus 11.9% of the MR-0 group. However, perhaps attributed to the greater response variation, a stronger outcome was that the students in the MR-2 + group were more likely to be in

---

**Table 1**

<table>
<thead>
<tr>
<th>Honesty item</th>
<th>Mischievous Response (MR) Group</th>
<th>(\chi^2(2, 1826))</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self: Endorsed “all” or “most” questions(^1)</td>
<td>MR-0 94.8(\text{a})</td>
<td>MR-1 90.1(\text{b})</td>
<td>MR-2 + 69.4(\text{c})</td>
</tr>
<tr>
<td>Peers: Endorsed “all” or “most” students(^2)</td>
<td>MR-0 48.5%</td>
<td>MR-1 46.2%</td>
<td>MR-2 + 42.9%</td>
</tr>
</tbody>
</table>

**Note.** Differing row subscript letters denote that the proportions for the mischievous response groups differed significantly at the .05 level.

\(^1\) How many questions in this survey did YOU answer honestly?  \(^2\) How many OTHER STUDENTS at your school do you think answered the questions in this survey honestly?

\(* * p < .001\)
Table 2
Means, Standard Deviations, F values, and Effect Sizes, by Mischievous Response (MR) Group

<table>
<thead>
<tr>
<th>Comparison indicators</th>
<th>MR index = 0</th>
<th>MR index = 1</th>
<th>MR index = 2+</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 1,524)</td>
<td>(n = 253)</td>
<td>(n = 49)</td>
<td></td>
</tr>
<tr>
<td><strong>Quality of life indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School connectedness</td>
<td>5–30</td>
<td>23.28 a, 4.70</td>
<td>22.68 a, 5.16</td>
<td>20.55 b, 8.33</td>
</tr>
<tr>
<td>SEHS-S covitality</td>
<td>36–150</td>
<td>112.5 a, 15.95</td>
<td>110.87 a, 18.57</td>
<td>105.27 b, 26.71</td>
</tr>
<tr>
<td>Survey orientation</td>
<td>3–18</td>
<td>11.94 a, 3.28</td>
<td>11.83 a, 3.94</td>
<td>10.14 b, 5.59</td>
</tr>
<tr>
<td><strong>CMH dual-factor measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMSLSS</td>
<td>5–25</td>
<td>20.95 a, 3.40</td>
<td>20.32 a, 3.75</td>
<td>19.22 b, 5.34</td>
</tr>
<tr>
<td>SDQ Emotion/Withdrawal</td>
<td>0–10</td>
<td>2.86 a, 2.42</td>
<td>3.21 b, 2.57</td>
<td>4.24 b, 3.61</td>
</tr>
<tr>
<td>SDQ Behavior/Conduct</td>
<td>0–10</td>
<td>2.18 a, 2.17</td>
<td>2.59 a, 2.37</td>
<td>3.88 b, 3.40</td>
</tr>
</tbody>
</table>

Note. MR = Mischievous Response index (range 0–7); SEHS-S = Social Emotional Health Survey – Secondary; CMH = complete mental health; BMSLSS = Brief Multidimensional Students’ Life Satisfaction Scale; SDQ = Strengths and Difficulties Questionnaire. *F* degrees of freedom for all tests = (2, 1823). Post hoc tests (Tukey’s) showed that the MR-0, MR-1, and MR-2 + means with different subscripts were significantly different (*p* < .01). *a* Adjusted *p* level needed for significance = .007.
the somewhat counterintuitive Vulnerable and Symptomatic but Content groups (22.4%) than were the MR-0 (7.8%) and MR-1 (8.3%) students. MR-0 students were more likely to fit within the Complete Mental Health classification.

### Discussion

During the secondary school years, youths assume a more central role in school-based mental health screening as they lend their personal perspective and voice to express the balance of strengths and distress that they might be experiencing during these critical developmental years (Weist & Baker-Sinclair, 1997). For this reason, it is recommended that universal school-based mental health screening include a validated self-report component (Levitt et al., 2007), which is particularly important for internalizing symptoms (Logan & King, 2002; Papandrea & Winefield, 2011). This increased reliance on student self-reports imparts an added responsibility on educators and school-based mental health professionals to carefully assess the veracity and authenticity of information derived from universal mental health screening. With respect to evaluating quality of self-reports in an actual universal screening at a comprehensive high school, the most robust finding of the present study was the validation that some students’ response profiles were consistent with a mischievous response pattern. The ways in which the MR-2 + students differed from their peers was consistent with Robinson-Cimpian’s (2014) finding that as few as two endorsed items is sufficient to identify mischievous responders. These students differed from the MR-0 and MR-1 groups on all comparisons and were more likely to display a Symptomatic but Content profile and less likely to display a Complete Mental Health profile. They also were less likely to acknowledge that they answered all or most items honestly. This combination of information about students with mischievous response patterns suggests that they may not be responding accurately. When conducting schoolwide screening, the responses of students with two or more mischievous responses should be considered with caution and the trustworthiness of their dual-factor classification scrutinized carefully.

For the MR-1 students, we acknowledge that some of the “mischievous” responses were legitimate. A student could have had more than 10 siblings (as does one of the authors) or might have truly preferred rum raisin ice cream (31 students indicated so). In addition, other factors might increase the odds of a student endorsing a single extreme response option (e.g., a child experiencing low economic circumstances would be more likely to not have access to regular dental visits). Nonetheless, the results showed that endorsing any single mischievous

<table>
<thead>
<tr>
<th>Distress criteria (rows)</th>
<th>Average-Low Distress</th>
<th>Average-High Distress</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDQ Emotional/Withdrawal and Behavior/Conduct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-score ≤ 60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mischievous Response Index = 0</td>
<td>15.6% (238)</td>
<td>64.7% (986)</td>
</tr>
<tr>
<td>Mischievous Response Index = 1</td>
<td>19.4% (49)</td>
<td>55.7% (141)</td>
</tr>
<tr>
<td>Mischievous Response Index = 2+</td>
<td>18.4% (9)</td>
<td>40.8% (20)</td>
</tr>
<tr>
<td>High Distress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ Emotional/Withdrawal and/or Behavior/Conduct</td>
<td>Troubled</td>
<td>Symptomatic but Content</td>
</tr>
<tr>
<td>T-score &gt; 60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mischievous Response Index = 0</td>
<td>11.9% (181)</td>
<td>7.8% (119)</td>
</tr>
<tr>
<td>Mischievous Response Index = 1</td>
<td>16.6% (42)</td>
<td>8.3% (21)</td>
</tr>
<tr>
<td>Mischievous Response Index = 2+</td>
<td>18.4% (9)</td>
<td>22.4% (11)</td>
</tr>
</tbody>
</table>

Note. SDQ = Strengths and Difficulties Questionnaire; BMSLSS = Brief Multidimensional Students’ Life Satisfaction Scale; SWB = subjective well-being. Each subscript letter denotes a subset of mischievous response group categories whose column proportions do not differ significantly from each other at the .05 level.

$$\chi^2 = 26.07 \, (6, \, 1826), \, p < .001, \, \text{Cramer’s } V = .084.$$
item, for whatever reason, was not systematically associated with meaningful mean differences when compared with students who endorsed zero mischievous items (MR-0) except that the MR-1 students were slightly more likely to be placed in the Troubled category and to report slightly less complete mental health.

Implications for Practice and Research

The findings of this study have some implications for the future refinement and use of the dual-factor approach for school-based mental health screening. When considered at the student classification level, our results indicated that only about 2% of our sample comprised obvious mischievous responders. Mischievous responders were identified in all four of the complete mental health categories; however, they were disproportionately placed into the Symptomatic but Content and the Troubled classifications. Within the complete mental health screening framework, students in the Troubled category (i.e., those with high distress and low subjective well-being) are considered to be at the highest level of risk and are the primary target for additional follow-up assessment and consideration. As resources allow, students in the Vulnerable category (i.e., those with average-low distress and low subjective well-being) are also viewed as potential candidates for support services aimed at bolstering their social-emotional health. The finding of disproportionate placement of mischievous responders among groups is of practical importance to care coordination teams because in the present study’s sample only a few (n = 9) of the Troubled category students (a primary target group of screening) had a mischievous response profile. Although school care coordination screening teams will want to be mindful of mischievous responses for this highest triage response group, in the present sample 78% of all students in the Troubled classification gave no indication of mischievous responding. Hence, we conclude that this preliminary evidence indicates that although mischievous responding occurred in the present study’s sample of students completing a universal mental health screening survey, the greatest number of students who presented with Vulnerable or Troubled profiles (the two main groups of interest to school care coordination teams) comprised students who appeared to provide meaningful self-report responses. This finding provided evidence that increases confidence about the potential usefulness of school-based mental health screening for most students.

The fact that surveys can easily include mischievous response items further supports their use within school-based mental health screening. We note that the specific questions used in this study’s mischievous response index adhered to Robinson-Cimpian’s (2014) recommendations; however, there are a multitude of other items that could be used to detect mischievous responding. Future studies should examine whether alternative items might more effectively identify mischievous responding and how differing school contexts and students’ sociocultural experiences might be related to response patterns.

This study also followed recommendations to incorporate a self-report question regarding the students’ opinions if the data could be trusted (i.e., How many questions in this survey did YOU answer honestly?), along with a cursory examination of response times as an index of careless responding (Meade & Craig, 2012). As the practice of school-based screening for mental health is on the rise (Bruhn et al., 2012), it will be important for those interested in conducting screenings via self-report to engage in practices that increase the likelihood of yielding high-quality data. Increased levels of careless responding have been shown to be associated with a variety of factors including a low level of interest and engagement in the survey, longer surveys, little communication between the respondent and the person(s) conducting the survey (e.g., online anonymity), and higher levels of environmental distractions (Meade & Craig, 2012). As the ease of online data collection grows with improved technology infrastructure, it will be important to keep screening surveys brief, in nondisturbing/proctored environments, and include content that is applicable and engaging to all students. Additionally, it will be important to communicate with students about the importance of the information obtained and how the nonanonymous results will be used to support students in need.
When conducting universal school-based screening, it is important to remember the goal of considering all student responses so that all students within a given population have the opportunity to be identified and referred for additional services. How, then, should practitioners and researchers proceed with handling student responses that may be mischievous? In many contexts, data that are considered problematic may be discarded; however, this approach of removing the data (or people) from further consideration is not advised within a screening context as the students with these response patterns may still be in need of additional services. Rather than merely excluding responses that fit the mischievous response pattern or removing such data in analyses, multiple gating screening approaches could be employed with the goal of successively and accurately narrowing down the population to identify those with true risk. Within a multiple gating screening approach, school staff follow up with a subset of the student population who are classified as at-risk after an initial screening, in order address any errors that may have occurred in the screening process (e.g., false positives; Walker, Small, Severson, Seeley, & Feil, 2014). Similarly, students who exhibit mischievous response patterns on an initial screening could be provided a more comprehensive assessment (e.g., additional items to assess the veracity of responses, interview, or gathering information from a different informant) in order to better discern whether the mental health data gathered from such students is a valid reflection of their true mental health state. Given the tendency for mischievous responders to rate themselves as being at higher risk and to self-report as being members of minority populations, a multiple gating approach might be able to reduce the error rate in screening, particularly false positives, and the overrepresentation of such populations. However, given that this is the first study of its kind, further use and investigation of mischievous responders, and the potential use of mischievous response items within a multiple gating approach, will help guide recommendations for school-based practice. Considering that mischievous response items can be easily incorporated into a first gate screening, inclusion in the first gate of a multiple gating screening procedure may be useful for identifying true positives prior to making intervention related decisions.

**Limitations**

This study used data derived from a survey implemented universally at one comprehensive high school. As such, the results reported herein do not generalize to all other school contexts or to screeners with different item content. It is possible that surveys with proportionally greater focus on positive student characteristics (e.g., social–emotional strengths), as in the present study’s survey, do not evoke mischievous responding in the same way that might be evoked by surveys that focus more heavily on risk behaviors (e.g., binge drinking). Similarly, this study’s survey was not anonymous. The fact that the students knew that their survey responses would be known to school officials might have decreased mischievous intentional- ity. Using identified rather than anonymous responding is associated with a decrease in careless responding (Meade & Craig, 2012) and may also be associated with a decrease in mischievous responding. Hence, the findings of this study do not generalize to more frequently used anonymous school surveys that focus on risk behaviors (e.g., Youth Risk Behavior Surveillance Survey; (CDC, 2015). Nonetheless, any effective and responsive schoolwide mental health screening will employ respectful, confidential, nonanonymous procedures similar to those employed in the present study. Furthermore, this study could have benefitted from the use of other measures, particularly considering the inconsistent psychometric characteristics with the SDQ (Ruchkin et al., 2008; Stevanovic et al., 2015) and the shortened version used in this study. There could be other options for instruments that are more viable in terms of validity, length, and cost-effectiveness. Also, although individual items within the survey were randomly ordered, surveys were given in the same order for all respondents. Order effects were not examined in this study; future studies should strive to ascertain whether response patterns are related to the order in which instruments are given. Finally, the study’s data were collected as part of an actual universal schoolwide screening carried out in partnership with a
comprehensive school, and as such, it has high context and clinical validity.

The results of this study cannot be interpreted to mean that all the MR-0 students provided authentic responses, but, rather, that they had lower odds of presenting a mischievous response pattern. Some MR-0 students might have had other factors that influenced their responses, such as satisficing (Barge & Gehlbach, 2012) and quick responses without careful consideration. Future research is needed to explore and contrast how other factors might influence or bias students’ responses to mental health screeners, such as presentation format (e.g., anchoring and response formats; Gehlbach & Barge, 2012). Mischievous responding is not the only possible influence on students’ responses. Instead of, or in addition to, responding mischievously, or “outright faking,” students may respond in a careless or random fashion, which can be even more difficult to detect, particularly considering that they are more likely to intermittently respond inattentively while taking the survey (Meade & Craig, 2012). Future research needs to examine the unique and shared influences of the various factors that influence student response styles. Additionally, although this study found significant differences between those who endorsed zero or one mischievous response item and those who endorsed two or more, future studies should examine whether endorsement of two items is the ideal cutpoint and/or if there are other significant differences between MR-2 responders and those who have even higher mischievous response rates.

With respect to additional limitations, we note that this study did not examine the optimal cutpoints used to define high and low values on the dual-factor distress and wellness continua (Kim, Furlong, Ng, & Huebner, in press), nor if an empirical algorithm (e.g., groups defined by latent class analysis; Rebelez, 2015) might increase the precision of categories used to organize the results of universal mental health screening. In addition, the number of students that opted out of the survey (n = 55) exceeded the number of mischievous responders (n = 49). We were unable to obtain information about this group of students. The efforts of school care teams will be enhanced by future research that attempts to better understand this subset of students. Finally, the present study did not address the most important aspect of universal screening, which is if the identified students later had access to available and effective services (Husky, Sheridan, McGuire, & Olfson, 2011; Reddy, Newman, De Thomas, & Chun, 2009).

Conclusion

To the extent that a rate of about 2% obvious mischievous responders is typical of what might be found in other samples, the raw number of students to be scrutinized, even in a comprehensive secondary school, is within the capacity of care coordination teams to manage. Hence, despite mean score differences on selected quality of life latent traits and differences in survey orientation, mischievous response patterns in the present study were minimally associated with overall trustworthiness of complete mental health screening when considered at the student classification level. The results provide support for universal school-based complete mental health screening by showing that a large majority of students in this sample provided responses that evidenced reasonably honest, authentic response profiles.

In closing, we note an interesting caveat. Despite evidence that most students provided what appear to be reasonably authentic responses, less than one half of the students perceived that their peers completed the survey honestly. Including response validity checks in school-based mental health survey, such as mischievous responding, could also be part of a social norming communication effort that aims to convey the value and utility of universal screening. To the extent that students, parents, staff, and school boards recognize and believe that most students are not haphazardly responding, they might consider the results more seriously, value them more, and be motivated to take action in support of students’ mental health needs.

References


tal health: Toward a more comprehensive understanding of youth functioning. American Journal of Orthopsychiatry, 80, 462–472. http://dx.doi.org/10.1111/j.1939-0025.2010.01049.x


This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.


Received January 19, 2016
Revision received April 27, 2016
Accepted April 30, 2016