

MELSON ET AL. (2016) ALCOHOLISM: CLINICAL AND EXPERIMENTAL RESEARCH

1 Title: Self-other Differences in Student Drinking Norms Research: The Role of Impression
2 Management, Self-deception and Measurement Methodology

3

4 Authors: Ambrose J Melson^a, Ph.D

5 Mental Health & Wellbeing, University of Glasgow, Academic Centre, Gartnavel Royal Hospital,
6 1055 Great Western Road, Glasgow, G12 0XH, United Kingdom

7

8 Rebecca Louise Monk, Ph.D

9 Derek Heim, Ph.D

10 Department of Psychology, Edge Hill University, Ormskirk, Lancashire, L39 4QP, United Kingdom.

11

12 Original submission: 27/04/16

13 Resubmission 1: 19/08/16

14 Resubmission 2: 12/09/16

15 Source of support: None

16

17

18

19

20

21

22

23 ^aCorrespondence concerning this article should be addressed to: Ambrose J Melson, Mental Health &
24 Wellbeing, University of Glasgow, Academic Centre, Gartnavel Royal Hospital, 1055 Great Western
25 Road, Glasgow, G12 0XH, United Kingdom ambrose.melson@glasgow.ac.uk

26 ABSTRACT

27

28 **Background:** Data-driven student drinking norms interventions are based on reported normative
29 overestimation of the extent and approval of an average student's drinking. Self-reported differences
30 between personal and perceived normative drinking behaviors and attitudes are taken at face value as
31 evidence of actual levels of overestimation. This study investigates whether commonly used data
32 collection methods and socially desirable responding may inadvertently impede establishing 'objective'
33 drinking norms.

34

35 **Methods:** UK students [N=421; 69% female; Mean age 20.22 years (SD = 2.5)] were randomly
36 assigned to one of three versions of a drinking norms questionnaire: The standard multi-target
37 questionnaire assessed respondents' drinking attitudes and behaviors (frequency of consumption, heavy
38 drinking, units on a typical occasion) as well as drinking attitudes and behaviors for an 'average
39 student'. Two deconstructed versions of this questionnaire assessed identical behaviors and attitudes
40 for participants themselves *or* an 'average student'. The Balanced Inventory of Desirable Responding
41 was also administered.

42

43 **Results:** Students who answered questions about themselves and peers reported more extreme
44 perceived drinking attitudes for the average student compared with those reporting solely on the
45 'average student'. Personal and perceived reports of drinking behaviors did not differ between multi-
46 and single-target versions of the questionnaire. Among those who completed the multi-target
47 questionnaire, after controlling for demographics and weekly drinking, socially desirable responding
48 was related positively with the magnitude of difference between students' own reported
49 behaviors/attitudes and those perceived for the average student.

50 **Conclusions:** Standard methodological practices and socially desirable responding may be sources of
51 bias in peer norm overestimation research.

52

53

INTRODUCTION

54

55 Drinking norms interventions are widely used in efforts to curb risky drinking practices among
56 students (DeJong et al., 2006, DeJong et al., 2009, Haines et al., 2005, Moore et al., 2013, Moreira et
57 al., 2009, Perkins, 2002, Perkins, 2003, Wechsler et al., 2003, Foxcroft et al., 2015). In these data-
58 driven intervention approaches, self-report questionnaires gauging personal drinking behaviors and
59 attitudes alongside matching perceptions for a normative target (e.g. the ‘average student’) feature
60 heavily throughout assessment, intervention activities and evaluations. Self-other differences or
61 differences between actual and perceived alcohol-related behaviors (i.e. *descriptive*) and
62 attitudes/approval (i.e. *injunctive*) support intervention efforts designed to counter errors in perception
63 by revealing 'actual' rather than attributed norms¹. This is often achieved by feeding back and
64 contrasting aggregate personal and peer norms to the target population in order to encourage revision
65 of perceptions and behavior in line with (typically) lower actual drinking norms. Similar data
66 collection exercises can be used to evaluate the impact of the intervention on behavior and perception
67 and update future iterations in long running intervention programs.

68

69 Evidence that the standard data collection methodologies used in this field provide a reliable and
70 accurate picture of young people’s actual and perceived drinking environments tends to be drawn from
71 the broader alcohol epidemiologic field (e.g. Babor et al., 2000, Midanik, 1988, Del Boca and Darkes,
72 2003). Investigations of data collection methodologies employed within the drinking norms field are

¹ Descriptive and injunctive norms are appropriate social psychological terms widely used in the student drinking norms literature to distinguish normative behavior from normative approval of behavior. However, we refer to drinking behavior and drinking attitudes throughout much of the paper because our focus is often individual-reporting behavior rather than normative processes.

73 rare (Pape, 2012). Accumulating evidence, however, suggests there may be elements of drinking
74 norms research methods that are potential sources of bias that may exaggerate peer norm
75 overestimation. One selective review of the research concluded that peer norm overestimation may be
76 an exaggerated phenomenon, with potential sources of exaggeration including sampling bias, a
77 tendency to use forced choice response options and limited attention to potential underestimation of
78 peer norms (Pape, 2012). The results of experimental studies have also questioned the ‘objectivity’ of
79 data obtained in the context of drinking norms research. In these studies, students have been found to
80 adjust their own reports downwards when exposed to information about their peer group’s standing
81 (Cunningham and Wong, 2013, Klein and Kunda, 1993, Lombardi and Choplin, 2010).

82

83 Melson and colleagues (2011, 2012) investigated whether the common methodological practice in
84 drinking norms research of questioning students conjointly about themselves and their peers impacted
85 response patterns underpinning peer norm overestimation. Here, school pupils who completed a
86 standard multi-target version of a drinking norms questionnaire, which included both personal and
87 peer (the ‘typical pupil’) alcohol-related measures, reported more extreme perceived peer attitudes and
88 were more likely to report that peers would consume alcoholic drinks compared to when the peer
89 target was assessed in isolation. The impact of the manipulation was limited to shifts in perception
90 responses rather than pupils’ own reported behavior and attitudes and key frequency of consumption
91 and drunkenness responses were unaffected, suggesting limitations to the generalizability of this effect.
92 As Melson et al. (2011) conducted their research with school pupils the extent to which these findings
93 may be observed in university/college student populations has not yet been examined. This is
94 important for a number of reasons. First, most drinking norms research and interventions have targeted
95 university and college students. University and college students’ age, developmental phase and likely
96 experience with alcohol means they may have more established drinking patterns and robust beliefs
97 about their wider peer group (e.g. Monk and Heim, 2016). As a result they may be less sensitive to

98 features of the measurement tool. It is also important to establish that Melson et al.'s (2011) findings
99 are not due to the demand characteristics of asking an adolescent population to complete sensitive
100 measures in a classroom environment (McCambridge and Strang, 2006, Percy et al., 2005). Recent
101 challenges reproducing evidence in psychological science suggest further caution against overreliance
102 on effects obtained from single studies (Open Science Collaboration, 2015).

103
104 Moreover, it is unclear whether the measurement effect reported by Melson and colleagues (2011)
105 indicates a specific methodological artifact or a broader motive of presenting an overly positive
106 version of oneself. Socially desirable responding (SDR) is the tendency to offer overly positive self-
107 descriptions during self-report questionnaire assessments (Paulhus, 1984, Paulhus, 2002).
108 Contemporary conceptualisations of SDR emphasize two dimensions: 'Impression management',
109 reflecting conscious regulation of personal characteristics, attributes and behaviors, so as to cast
110 oneself in a favorable light; and 'self-deception', an unconscious propensity to think about oneself in
111 overly positive self-esteem maintaining ways when retrieving information during response (Paulhus,
112 2002, Paulhus, 1984, Holtgraves, 2004). In this vein, North American College students scoring highly
113 on a measure of impression management reported consuming 33% fewer drinks and lower AUDIT
114 scores than those in the moderate or low range. Self-deception was unrelated to drinking behavior but
115 was associated with reporting fewer alcohol-related problems (Davis et al., 2010, Lanyon and Carle,
116 2007, Paulhus, 1991, Paulhus, 2002, Paulhus and Reynolds, 1995, Paulhus et al., 1995). These
117 findings suggest that students may intentionally distort reports of their drinking behavior and related
118 problems so as to cast themselves favorably to others. In limited circumstances they are also prone to
119 the effects of self-deception.

120
121 Recent research has highlighted SDR as a potential source of bias in self-reported student drinking
122 responses (Davis et al., 2010). As a result SDR holds promise for understanding whether the self-other

123 differences observed in student drinking norms research reliably indicate drinking norm discrepancies
124 or, to some extent, a socially motivated drive to present an overly positive version of oneself. Different
125 dimensions of SDR may also be important for determining the nature of self-other difference
126 measurement. Specifically, if self-other differences for peer group drinking behavior and attitudes are
127 associated with impression management, then self-other differences may reflect intentional strategies
128 to present an unrealistic and overly positive version of oneself. Associations with self-deception, on
129 the other hand, may point towards a role for unconscious self-esteem maintaining biases in self-other
130 difference reporting. Investigating the role of SDR in reported self-other differences will help elucidate
131 the extent to which SDR poses a risk to reliable and valid measurement of self-other differences.
132 Clarifying the likely motivational base of any SDR bias will also inform our understanding of the
133 underlying processes involved and guide the development of effective strategies to minimize socially
134 desirable response patterns.

135

136 Given their importance within student drinking norms research and interventions, remarkably few
137 studies have investigated the reliability and validity of ‘self’ and ‘other’ drinking responses. The
138 present study seeks to address this shortcoming in the literature by extending one of few
139 methodological studies in this field (Melson et al., 2011). Consistent with earlier research, it was
140 predicted that responses to a drinking norm questionnaire, which assesses personal and perceived
141 drinking-related behaviors or attitudes conjointly, will differ from those that assess personal and
142 perceived measures in isolation. Although observing the presence of specific measurement artifacts is
143 important, understanding when and why they are likely to arise is crucial for advancing knowledge
144 that may support effective strategies to limit threats to objective measurements of drinking norms.
145 Therefore the central aim of this research was to investigate whether SDR plays a role in self-other
146 differences reported by students, with the expectation that SDR would be positively associated with
147 self-other differences.

148

149

150 Participants and Procedures

151 The research took place at two UK universities. Students (18-30 years) who had consumed alcohol in
152 the past year were invited to participate in an online ‘Student Drinking Survey’ via a URL advertised
153 on university social media, communication networks and research participation pools. Following self-
154 selection, eligible participants provided informed consent prior to allocation to one of three different
155 study conditions via a randomization function embedded in the online questionnaire. Ethics review
156 committees approved the research at both institutions.

157

158 Design

159 Between-participants experimental design, with randomization to one of three different versions of a
160 drinking norms questionnaire [(i) multi-target (MT) version or (ii) single-target ‘personal’ (ST^{PERS}) or
161 (iii) single-target ‘average student’ (ST^{AS}) versions].

162

163 Measures

164 In overview, students provided demographic details (age, gender, ethnicity, year of study) and
165 completed target-specific (‘personal’ and/or perceived ‘average student’) measures of drinking
166 behavior and attitudes, as well as a measure of socially desirable responding.

167

168 Three different versions of a questionnaire were constructed. The first version was designed to closely
169 resemble a standard type of questionnaire employed in this field (e.g., Haines et al., 2005) and
170 included items to record students’ own alcohol-related behaviors and attitudes, in addition to their
171 perceptions of each behavior and attitudes for the ‘average student your own age at your university’
172 [i.e., a multi-target (MT) version]. Two single-target (ST) questionnaires included items corresponding

173 to a single target [i.e. single-target: ‘personal’ (ST^{PERS}) or ‘average student’ (ST^{AS})]. Notwithstanding
174 the omission of the other target, both ST versions of the questionnaire were identical to the full MT
175 version.

176

177 *Drinking behavior*

178 Drinking behavior and perceptions of other students’ behavior were assessed using original and
179 modified AUDIT-C items (Bush et al., 1998). Referring to the past 12 months, students reported how
180 often a drink containing alcohol had been consumed, how many units of alcohol were drunk on a
181 typical occasion and how frequently eight units of alcohol were consumed on one occasion. While the
182 original versions of the AUDIT-C items were used to assess students’ own behavior, modified items
183 assessed perceptions of identical behaviors for the ‘average student your own age at the university’.
184 Original and modified versions of these items differed only in the specific target-referent of the item
185 (e.g. ‘How often do *you* [the average student your own age at the university] have a drink containing
186 alcohol?’). Responses were used to calculate mean monthly frequencies of consumption, heavy
187 drinking and the usual quantity of units consumed on drinking occasions. Recent seven-day
188 consumption was obtained by asking students to record the units of alcohol consumed on each of
189 seven days for a ‘typical week in the past month’. Accompanying information provided guidance on
190 the UK alcohol unit content of popular drinks and volumes.

191

192 *Drinking attitudes*

193 Attitudes and perceived attitudes to drinking were assessed using a scale described by Lewis et al.
194 (2010) in a study with US college students. After removing four items of limited cultural relevance in
195 a UK student population, the acceptability of 11 different alcohol-related behaviors were rated by
196 students for themselves and/or the acceptability of each behavior perceived for the average student the
197 respondent’s own age. Representative items include ‘Playing drinking games’ and ‘Drinking alcohol

198 daily'. Ratings were scored as unacceptable (1) through to acceptable (7) and summed to create
199 separate personal and average student attitude indexes (Min=11, Max=77). Internal consistency of
200 each index was good ($\alpha=0.81-0.82$) (Paulhus et al., 1995).

201

202 *Socially desirable responding*

203 Socially desirable responding was assessed using the Balanced Inventory of Desirable Responding-
204 Version 6 (BIDR: Paulhus, 1991), a 40-item measure of the propensity to provide overly positive self-
205 descriptions. Twenty positive and negative items tap the self-deception dimension of SDR, reflecting
206 unintended enhancement of personal abilities and qualities, or denial of undesirable ones. Twenty
207 positive and negative items also measure impression management, reflecting purposeful inflation of
208 socially desirable behaviors and qualities. Representative items are 'I never regret my decisions' (self-
209 deception) and 'I never cover up my mistakes' (impression management), with all items recorded on
210 7-point scales (*not true to very true*). The BIDR has received support as a valid measure of providing
211 overly positive self-descriptions via convergent and discriminant associations with other measures of
212 SDR and exaggerated virtue, adjustment, coping, self-esteem, 'lie' scales and under different role play
213 instructions (Lanyon and Carle, 2007, Paulhus, 1991, Paulhus et al., 1995). Importantly, overly
214 positive self-descriptions are only reflected in endorsement of extreme response categories 6-7 or 1-2
215 (depending on positive or negative framing of items), indicating responses which are too good to be
216 true. These are scored as '1' with all other response categories scored '0' (Paulhus, 1991). Impression
217 management and self-deception item scores are summed separately (Min = 0, Max = 20 for both). In
218 the present study internal consistency of both subscales was adequate ($\alpha = .67-.76$) and within the
219 normal range (e.g., Davis et al., 2010, Paulhus, 1991).

220

221 **Analysis**

222 Data from the two institutions were pooled to create a single dataset. We considered a nested approach
223 to analyses to account for possible within-cluster similarity of responses at the two institutions.
224 However, intraclass correlation coefficients for drinking behaviors, attitudes and SDR were very small
225 and indicated that 1% or less of variability was accounted for at the institution-level. Given limited
226 evidence that responses were clustered within institutions we analyzed the pooled dataset. Frequency
227 of heavy drinking and seven-day consumption distributions were positively skewed and values were
228 log transformed for main analyses, although descriptive statistics retain the untransformed values for
229 ease of interpretation. To address the first hypothesis, personal and perceived versions of the AUDIT-
230 C items and attitude indexes were compared between the MT and corresponding ST versions of the
231 questionnaire using *t*-tests and Cohen's *d*. For the second hypothesis, analyses were necessarily
232 restricted to students who completed the MT version of the questionnaire. Relationships between study
233 variables were initially examined using zero-order correlations. Consistent with other studies, self-
234 other difference values were calculated by subtracting personal from perceived response values for
235 each AUDIT-C item and the drinking attitude indexes (Carey et al., 2006). Four hierarchical regression
236 models were then constructed to examine whether SDR accounted for unique variance in self-other
237 differences beyond demographics and recent seven-day consumption. Self-other differences were used
238 as the key outcome variables in regression analyses based on their theoretical and practical
239 significance in normative drinking research and interventions.

240

241

RESULTS

242

243 A total of 421 participants across the two university institutions provided complete data, 322 (76%)
244 from one institution and 99 (24%) from a second. As participants self-selected into the study,
245 participation rates are not available. These samples were similar to the relevant student rolls on the
246 proportion reporting White ethnicity, but female gender, younger, and undergraduate students were

247 overrepresented. The pooled dataset was over two thirds (69%) female, primarily undergraduate (94%)
248 with an average age of 20.22 years (SD = 2.5). Half (50%) identified as ‘White British’ (or another
249 White UK nationality), 38% White other and 12% Mixed, Black or Asian. Following randomization,
250 142 students completed the MT version of the questionnaire, 158 completed the ST^{PERS} version and
251 121 completed the ST^{AS} version [$\chi^2_{(2, 421)} = 4.91, p > .05$]. Randomization of participants to complete
252 the different version of the questionnaire was successful. Participants’ did not differ significantly
253 across demographic variables: gender [$\chi^2_{(2, 421)} = 0.622, p > .05$], age [$F_{(2, 420)} = 0.22, p > .05$] year of
254 study [$\chi^2_{(2, 419)} = 0.1, p > .05$] or ethnicity [$\chi^2_{(2, 411)} = 0.98, p > .05$].

255

256 **Questionnaire version**

257 Table 1 presents mean (SD) drinking behavior responses and attitude index scores obtained from the
258 MT and corresponding ST versions of the questionnaire. Responses of students who completed the
259 MT and ST^{PERS} versions of the questionnaire did not differ for students’ own reported frequencies of
260 consumption, heavy drinking, quantity of units on a typical occasion or the attitude index. There were
261 also no differences between the MT and ST^{AS} versions of the questionnaire in reported perceptions of
262 the average student’s drinking behavior. However, students who completed the MT version of the
263 questionnaire had higher perceived attitude index scores than those who completed the ST^{AS} version.

264

265 TABLES 1 AND 2 AROUND HERE (APPENDED)

266

267 **Socially desirable responding**

268 Table 2 presents zero-order correlations for SDR subscales and key study variables. Inter-correlations
269 for students’ personal and perceived drinking behaviors and demographics were broadly in line with
270 the existing literature. Personal drinking behaviors and corresponding perceptions tended to be
271 positively associated, although personal attitudes were unrelated to perceptions of the average

272 student's attitudes. Increasing age was associated with lighter drinking and more conservative attitudes,
273 while male gender tended to be associated with heavier personal drinking and perceived drinking
274 frequencies. Impression management was positively associated with female gender, perceptions of the
275 average student's drinking frequency and negatively associated with students' own typical unit
276 consumption and attitudes. Self-deception was positively associated with perceptions of the average
277 student's heavy drinking. Two aspects of the relationship between SDR and AUDIT-C or attitudes
278 measures are noteworthy. First, SDR was associated with all three AUDIT-C drinking behaviors and
279 drinking attitudes, either via 'personal' or perceived 'average student' responses. Second, SDR was
280 positively associated with reported perceptions of the average student's behaviors and negatively
281 associated with students' own reported behaviors or attitudes.

282

283 TABLE 3 AROUND HERE (APPENDED)

284

285 Whether SDR accounted for unique variance in self-other differences was investigated by regressing
286 the four self-other difference values separately on demographics (step 1), seven-day consumption (step
287 2) and SDR subscales (step 3).

288

289 In each final model (Table 3), seven-day consumption and SDR subscales independently predicted
290 self-other differences. Lower seven-day consumption strongly predicted larger self-other differences in
291 each model (β 's = -0.32 to -0.60, $ps < .001$). Higher impression management predicted larger self-
292 other differences for frequency of consumption, typical units and attitudes (β 's = 0.19 to 0.22, ps
293 $< .05$) and higher self-deception predicted larger self-other differences for heavy drinking frequency (β
294 = 0.16, $p < .05$). Increasing age and female gender initially predicted larger self-other differences for
295 drinking frequency and attitudes, but were not independent of other factors in final models. Step 1 of
296 the drinking frequency and attitude index models accounted for 5% and 8% of the variance in self-

297 other differences respectively. Seven-day consumption accounted for an additional 11-30% of variance
298 across all four models, before SDR subscales accounted for a further 3-4% of the variance.

299

300 Possible moderation of the relationship between SDR subscales and self-other differences by gender
301 was investigated by entering the product of SDR subscales and gender as interaction terms. These did
302 not account for additional variance in self-other differences (data not reported). Figure 1 (a-d) depicts
303 the increasing magnitude of self-other differences across low, medium and high impression
304 management or self-deception scores (based on tertile splits) for AUDIT-C drinking behaviors and
305 attitudes.

306 FIGURE 1(a-d) AROUND HERE (APPENDED)

307

308 DISCUSSION

309

310 This study investigated whether commonly used data collection methods and socially desirable
311 responding (SDR) may contribute to frequently observed self-other differences within student drinking
312 norms research. For the first time, we report that self-other differences appear to increase in relation to
313 SDR, a reporting bias possibly reflecting a tendency to provide overly positive self-descriptions. A
314 modest, but consistent, 3-4% of variance in self-other differences was accounted for by SDR across
315 three dimensions of student drinking behavior and drinking attitudes, raising the possibility that SDR
316 may be a common feature of observed self-other differences. Classifying participants as low, medium
317 or high on SDR indicates that students who tend to respond on a socially desirable basis report
318 markedly larger self-other differences than students less prone to SDR.

319

320 Whether SDR plays a causal role in determining the magnitude of self-other differences cannot be
321 determined from the cross-sectional design of this part of the study. Available evidence and the results

322 of the present study, however, lend support to the contention that self-other differences may be
323 exaggerated as a result of students' potentially casting themselves in too favorable a light. First, the
324 measure of SDR used in this study has received support as a valid measure of providing overly
325 positive self-descriptions, increasing confidence that desirable responding influences self-other
326 difference measurement (Lanyon and Carle, 2007, Paulhus, 1991, Paulhus et al., 1995). Second, the
327 question of whether observed self-other differences are objectively larger for high SDR scorers
328 because they tend to be more moderate drinkers has partly been addressed in previous research carried
329 out by Davis and colleagues (2010). Their research examined the extent to which differences in
330 drinking reports across levels of impression management were independent of traits that would predict
331 genuinely more moderate consumption. Consistent with an account based on biased reporting rather
332 than true differences in behavior, statistically adjusting for impulsivity-constraint, a key predictor of
333 drinking behavior (Hair and Hampson, 2006, Granö et al., 2004, Curcio and George, 2011) did not
334 alter the significantly lower levels of alcohol use, hazardous drinking and problems reported by
335 students scoring high on impression management. Furthermore, if the present findings were due to
336 SDR and self-other differences sharing a common etiology, or associations, with unmeasured 'third'
337 variables, we would expect SDR patterns to parallel known reporting patterns including the positive
338 association between students' own drinking and peer perceptions (Carey et al., 2006). However, the
339 present results indicate that SDR is both negatively correlated with students' own behavior and also
340 positively correlated with perceptions of peer behavior. The current findings therefore suggest that
341 socially desirable responding may exaggerate measured discrepancies between students' real and
342 perceived drinking patterns.

343

344 As well as casting a degree of doubt on the size of self-other differences frequently obtained in
345 drinking norms research, the current findings may have implications for intervention approaches.
346 Importantly, these implications do not speak directly to the efficacy of normative intervention

347 approaches, but may be most relevant when considering the design and conduct of interventions. Data-
348 driven norms interventions are unique in the field because self-other difference measurements are
349 typically used during assessment, intervention and evaluation. In extreme cases of desirable
350 responding, it is therefore conceivable that observed self-other differences may differ markedly from
351 those which are true of the population, potentially influencing practitioner decision-making concerning
352 where scarce resources are targeted. However, highlighting larger self-other differences as part of a
353 normative intervention may also be preferred by practitioners due to the heightened salience and
354 increased opportunity to alert students to perception-behavior discrepancies. In turn, students exposed
355 to larger self-other differences may experience heightened dissonance that motivates a reduction in
356 drinking; raising the possibility that SDR may actually increase the potency of an intervention.
357 However, it seems prudent to suggest that careful consideration be given to the pros and cons of
358 highlighting self-other differences that may, to a greater or lesser extent, reflect SDR. Although there
359 may be some advantage to including larger self-other differences as part of an intervention, promoting
360 accurate information is often considered a key element of normative intervention.

361

362 A further aim was to investigate whether commonly used data collection methods may contribute to
363 self-other differences. Here, students questioned about their own drinking behaviors and attitudes, as
364 well as those of the average student, reported more permissive attitudes among their peers than
365 students asked solely about their perceptions of the average student. In contrast, students' own
366 attitudes did not differ between questionnaires and reports of drinking behaviors (both personal and
367 perceived) and appeared to be more robust to the questionnaire manipulation. Underscoring the need
368 for research to improve measurement within drinking norms research (Monk and Heim, 2014, Pape,
369 2012, Simons-Morton and Kuntsche, 2012), these findings extend those of Melson et al. (2011) to
370 university settings which are frequently the target of drinking norms research and interventions. In
371 doing so, we confirm that there appears to be a reproducible effect of the type of questionnaire used to

372 measure perceived drinking attitudes, which has the potential to increase the magnitude of self-other
373 differences. Data in this study were collected remotely via an online survey, and used a different set of
374 measures than in previous research. This suggests the measurement effect is not in itself an artifact of a
375 specific set of attitude statements or of questioning a young sample in classroom environments
376 (McCambridge and Strang, 2006, Percy et al., 2005, Open Science Collaboration, 2015). Importantly,
377 this study also provides further evidence that student drinking behavior and perception responses
378 appear to be robust to multi- or single-target presentation. This increases confidence that self-other
379 differences/normative overestimation of drinking behavior may be unaffected by this aspect of
380 standard methodology. The multi-target questionnaire used in this research assessed a single order of
381 target presentation ('self-then-peer'), which does not control for possible ordering effects, but is
382 consistent with available guidance and likely to reflect applied practice (Haines et al., 2005). The
383 possible inflation of the true magnitude of discrepancy for personal and perceived drinking attitudes
384 means researchers may prefer to consider alternatives to multi-target questionnaires when attempting
385 to quantify accurately this type of discrepancy.

386

387 Our finding that SDR is related to self-other differences offers a useful starting point for investigating
388 possible mechanisms responsible for the differences obtained in perceived peer attitudes using multi-
389 and single-target questionnaires. However, SDR was unrelated to perceptions of the average student's
390 drinking attitudes, suggesting that alternative accounts may be needed to explain this effect. One
391 possibility is that multi-target questionnaires encourage a more context-specific form of SDR, distinct
392 from the stable trait operationalization of SDR in this study. Research that can manipulate social
393 desirability demands during assessments may help to identify whether self-other differences vary as a
394 function of more contextually-dependent social desirability (e.g., Holtgraves, 2004).

395

396 A strength of the study was the use of a multi- rather than single- dimensional measure of SDR
397 (Paulhus, 1991, Crowne and Marlowe, 1960), enabling tentative hypotheses about SDR processes that
398 may influence drinking norms reports. Self-deception was positively associated with frequency of
399 heavy drinking self-other differences, while impression management was positively related to self-
400 other differences for drinking frequency and units on a typical occasion and attitudes. Biased estimates
401 of heavy drinking self-other differences may therefore be obtained when students unintentionally
402 retrieve overly positive self-esteem maintaining information during questionnaire assessments. In
403 contrast, biased self-other differences for drinking frequency, quantity and attitudes may be obtained
404 because students consciously distort reports in order to present a more favorable version of themselves.
405 One possible explanation for the association of self-deception with frequency of heavy drinking self-
406 other differences may lie in the self-esteem maintaining function of self-deception. Self-deceptive
407 responding may become increasingly likely as questions deal with risky or health-compromising
408 behaviors such as heavy drinking. Consistent with this account, Davis et al. (2010) also found that self-
409 deception was unrelated to basic consumption reports but was negatively associated with reporting
410 drinking problems. Strategies to limit the impact of SDR may also benefit from consideration of the
411 different SDR processes. Thus, while confidentiality is often stressed during questionnaire assessments,
412 a procedure that may help to reduce the impact of impression management, this is unlikely to address
413 the hypothesised self-deception processes. Future research may therefore usefully develop and test
414 targeted messages designed to minimize the impact of self-deception, for instance by encouraging
415 more careful or balanced information retrieval.

416

417 The different SDR processes reported here also hold possible implications for applied drinking norms
418 interventions. Credibility of normative feedback is a key moderator of intervention effectiveness (e.g.,
419 Thombs et al., 2004, Polonec et al., 2006). Among students with strong impression management biases
420 (i.e., reflecting conscious and intentional distortion of responses to present a more favorable version of

421 themselves), projecting this response bias onto peers could lead some to dismiss normative feedback
422 components of interventions as poor reflections of real world norms. Additionally, among students
423 who self-deceive, a drinking norms intervention might actually help to counter self-deceptive
424 responding if, following exposure, respondents are encouraged to confront biased information retrieval.
425 These possible implications are speculative and require further examination in carefully designed
426 studies or by incorporating measures of SDR into evaluations of drinking norm interventions.

427

428 There are several possible limitations to this study. We used a self-selecting sample with an
429 overrepresentation of females and younger undergraduate students. Notwithstanding these limitations,
430 we note that the experimental findings reported closely resemble those obtained in previous research
431 using a gender-balanced cohort of much younger of adolescent pupils (Melson et al., 2011), increasing
432 confidence that our results are unlikely to be due to selection bias. The current research used distal
433 (average student) targets because this is the dominant target within this field of research. However, it
434 remains to be investigated whether perceptions of a more proximal target such as close friends would
435 be less sensitive to measurement effects or unrelated to SDR. This area of research therefore warrants
436 further exploration. Overall, while the experimental findings are consistent with earlier research, the
437 role of socially desirable responding is novel and further investigation of its relationship to self-other
438 differences and peer norm overestimation seems prudent.

439

440 In conclusion, self-reported differences between personal and perceived drinking behaviors and
441 attitudes are frequently taken at face value as evidence of actual levels of normative overestimation
442 within student populations. The present study indicates that reported self-other differences in drinking
443 behaviors and attitudes may partly reflect socially desirable responding and be a possible by-product
444 of standard methodological practices. Overestimation of peer drinking, as commonly reported, may not
445 reflect 'objective' levels of overestimation in student populations.

- 448 Babor T, Steinberg K, Anton R, Del Boca F (2000) Talk is cheap: Measuring drinking outcomes in
449 clinical trials. *J Stud Alcohol* 61:55-63.
- 450 Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA (1998) The AUDIT alcohol consumption
451 questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med*
452 158:1789.
- 453 Carey KB, Borsari B, Carey MP, Maisto SA (2006) Patterns and importance of self-other differences in
454 college drinking norms. *Psychol Addict Behav* 20:385-393.
- 455 Crowne DP, Marlowe D (1960) A new scale of social desirability independent of psychopathology. *J*
456 *Consult Psychol* 24:349-354.
- 457 Cunningham JA, Wong HTA (2013) Assessing the immediate impact of normative drinking information
458 using an immediate post-test randomized controlled design: Implications for normative
459 feedback interventions? *Addict Behav* 38.
- 460 Curcio AL, George AM (2011) Selected impulsivity facets with alcohol use/problems: the mediating
461 role of drinking motives. *Addict Behav* 36:959-964.
- 462 Davis CG, Thake J, Vilhena N (2010) Social desirability biases in self-reported alcohol consumption
463 and harms. *Addict Behav* 35:302-311.
- 464 DeJong W, Schneider SK, Towvim LG, Murphy MJ, Doerr EE, Simonsen NR, Mason KE, Scribner RA
465 (2006) A multisite randomized trial of social norms marketing campaigns to reduce college
466 student drinking. *J Stud Alcohol* 67:868-879.
- 467 DeJong W, Schneider SK, Towvim LG, Murphy MJ, Doerr EE, Simonsen NR, Mason KE, Scribner RA
468 (2009) A multisite randomized trial of social norms marketing campaigns to reduce college
469 student drinking: a replication failure. *Subst Abus* 30:127-149.
- 470 Del Boca FK, Darkes J (2003) The validity of self-reports of alcohol consumption: state of the science
471 and challenges for research. *Addiction* 98:1-12.
- 472 Foxcroft DR, Moreira MT, Almeida Santimano NM, Smith LA (2015) Social norms information for
473 alcohol misuse in university and college students. *The Cochrane Library*.
- 474 Granö N, Virtanen M, Vahtera J, Elovainio M, Kivimaki M (2004) Impulsivity as a predictor of smoking
475 and alcohol consumption. *Pers Individ Differ* 37:1693-1700.
- 476 Haines MP, Perkins HW, Rice RM, Barker G (2005) A guide to marketing social norms for health
477 promotion in schools and communities, in *Series A guide to marketing social norms for health*
478 *promotion in schools and communities*, National Social Norms Resource Center, Virginia.
- 479 Hair P, Hampson SE (2006) The role of impulsivity in predicting maladaptive behaviour among female
480 students. *Pers Individ Differ* 40:943-952.
- 481 Holtgraves T (2004) Social desirability and self-reports: Testing models of socially desirable
482 responding. *Pers Soc Psychol B* 30:161-172.
- 483 Klein WM, Kunda Z (1993) Maintaining self-serving social comparisons: Biased reconstruction of one's
484 past behaviors. *Pers Soc Psychol B* 19:732-739.
- 485 Lanyon RI, Carle AC (2007) Internal and external validity of scores on the Balanced Inventory of
486 Desirable Responding and the Paulhus Deception Scales. *Educational and Psychological*
487 *Measurement* 67:859-876.
- 488 Lewis MA, Neighbors C, Geisner IM, Lee CM, Kilmer JR, Atkins DC (2010) Examining the associations
489 among severity of injunctive drinking norms, alcohol consumption, and alcohol-related

490 negative consequences: The moderating roles of alcohol consumption and identity. *Psychol*
491 *Addict Behav* 24:177-189.

492 Lombardi MM, Choplin JM (2010) Anchoring and estimation of alcohol consumption: implications for
493 social norm interventions. *J Alcohol Drug Educ* 54:53-71.

494 McCambridge J, Strang J (2006) The reliability of drug use data collected in the classroom: what is the
495 problem, why does it matter and how should it be approached? *Drug Alcohol Rev* 25:413-418.

496 Melson AJ, Davies JB (2012) Interpreting questionnaire design effects: an update on Melson et
497 al.(2011). *Addiction* 107:1883-1884.

498 Melson AJ, Davies JB, Martinus T (2011) Overestimation of peer drinking: error of judgment or
499 methodological artefact? *Addiction* 106:1078-1084.

500 Midanik LT (1988) Validity of self-reported alcohol use: a literature review and assessment. *Brit J*
501 *Addict* 83:1019-1029.

502 Monk RL, Heim D (2014) A systematic review of the Alcohol norms literature: A focus on context.
503 *Drugs: education, prevention and policy* 21:263-282.

504 Monk RL, Heim D (2016) Alcohol-related expectancies in adults and adolescents: Similarities and
505 disparities. *Adicciones* 28:35-40.

506 Moore GF, Williams A, Moore L, Murphy S (2013) An exploratory cluster randomised trial of a
507 university halls of residence based social norms marketing campaign to reduce alcohol
508 consumption among 1st year students. *Substance Abuse Treatment, Prevention, and Policy*
509 8:15-15.

510 Moreira MT, Smith L, Foxcroft D (2009) Social norms interventions to reduce alcohol misuse in
511 university or college students, in *Series Social norms interventions to reduce alcohol misuse in*
512 *university or college students, Cochrane Database of Systematic Reviews* 3.

513 Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science*
514 349:944-951.

515 Pape H (2012) Young people's overestimation of peer substance use: an exaggerated phenomenon?
516 *Addiction* 107:878-884.

517 Paulhus D (1984) Two-component models of socially desirable responding. *Journal of personality and*
518 *social psychology* 46:598.

519 Paulhus D (1991) Measurement and control of response bias, in *Measures of personality and social*
520 *psychological attitudes, Vol. 1, Measures of personality and social psychological attitudes*
521 *(ROBINSON J, SHAVER P, WRIGHTSMAN L eds), pp 17-59, Academic Press, San Diego, CA.*

522 Paulhus D (2002) Socially desirable responding: The evolution of a construct, in *The role of constructs*
523 *in psychological and educational measurement, The role of constructs in psychological and*
524 *educational measurement (BRAUN HI, JACKSON DN, WILEY DE eds), pp 51-66, Lawrence*
525 *Erlbaum, Mahwah: New Jersey.*

526 Paulhus D, Reynolds S (1995) Enhancing target variance in personality impressions: highlighting the
527 person in person perception. *J Pers Soc Psychol* 69:1233-1242.

528 Paulhus DL, Bruce MN, Trapnell PD (1995) Effects of self-presentation strategies on personality
529 profiles and their structure. *Pers Soc Psychol B* 21:100-100.

530 Percy A, McAlister S, Higgins K, McCrystal P, Thornton M (2005) Response consistency in young
531 adolescents' drug use self-reports: a recanting rate analysis. *Addiction* 100:189-196.

532 Perkins HW (2002) Social norms and the prevention of alcohol misuse in collegiate contexts. *J Stud*
533 *Alcohol* 14:164-172.

534 Perkins HW (2003) *The social Norms approach to preventing school and college age substance abuse:*
535 *a handbook for educators, counselors, and clinicians, Jossey-Bass, San Francisco.*

536 Polonec LD, Major AM, Atwood LE (2006) Evaluating the believability and effectiveness of the social
537 norms message" most students drink 0 to 4 drinks when they party". Health communication
538 20:23-34.

539 Simons-Morton B, Kuntsche E (2012) Adolescent estimation of peer substance use: why it matters.
540 Addiction 107:885-886.

541 Thombs DL, Dotterer S, Olds RS, Sharp KE, Raub CG (2004) A close look at why one social norms
542 campaign did not reduce student drinking. J Am Coll Health 53:61-68.

543 Wechsler H, Nelson TE, Lee JE, Seibring M, Lewis C, Keeling RP (2003) Perception and reality: a
544 national evaluation of social norms marketing interventions to reduce college students' heavy
545 alcohol use. J Stud Alcohol 64:484-494.

Table 1. Mean (SD) drinking and attitude index scores for multi- and single-target versions of the questionnaire

	Questionnaire version				Questionnaire version			
	MT	ST ^{PERS}	<i>t</i>	<i>d</i>	MT	ST ^{AS}	<i>t</i>	<i>d</i>
AUDIT Frequency of consumption	7.38 (4.99)	7.11 (4.93)	0.48 ^{NS}	0.05	10.9 (3.96)	11.17 (3.97)	0.54 ^{NS}	0.07
AUDIT Typical occasion units	6.1 (3.06)	5.80 (2.83)	0.88 ^{NS}	0.1	7.37 (2.61)	7.28 (2.61)	0.29 ^{NS}	0.04
AUDIT heavy drinking	2.43 (3.81)	2.31 (3.65)	0.3 ^{NS}	0.03	4.54 (5.66)	4.05 (3.71)	0.09 ^{NS}	0.02
Drinking attitudes	48.51 (9.99)	46.66 (11.43)	1.48 ^{NS}	0.17	60.49 (8.65)	56.28 (10.26)	3.61 ^{***}	0.44

*** $P < .001$; ^{NS} $P > .05$; *d* = Cohen's *d*

MT: Multi-target version of the questionnaire; ST^{PERS}: Single-target 'personal' version of the questionnaire; ST^{AS}: Single-target 'average student' version of the questionnaire.

AUDIT frequency of heavy drinking values were log transformed

Table 3. Hierarchical regression analyses of socially desirable responding and self-other differences in frequency of consumption, typical occasion units, frequency of heavy drinking and attitudes

AUDIT frequency of consumption SOD						AUDIT typical occasion units SOD					
	Variable	R^2	ΔR^2	beta	t		Variable	R^2	ΔR^2	beta	t
Step 1	Age	.05	.05*	0.20	2.36*		Age	.03	.03 ^{NS}	0.11	1.28 ^{NS}
	Gender			0.10	1.2 ^{NS}		Gender			0.14	1.67 ^{NS}
Step 2	Age	.28	.23***	0.05	0.62 ^{NS}		Age	.13	.10***	0.01	0.00 ^{NS}
	Gender			-0.02	-0.21 ^{NS}		Gender			0.06	0.63 ^{NS}
	7-day consum			-0.52	-6.65***		7-day consum			-0.34	-4.00***
Step 3	Age	.32	.04*	0.04	0.58 ^{NS}		Age	.17	.04*	0.01	0.14 ^{NS}
	Gender			-0.05	-0.66 ^{NS}		Gender			0.03	0.42 ^{NS}
	7-day consum			-0.5	-6.55***		7-day consum			-0.32	-3.78***
	BIDR SD			0.00	0.06 ^{NS}		BIDR SD			-0.05	-0.65 ^{NS}
	BIDR IM			0.21	2.83**		BIDR IM			0.22	2.63*
AUDIT frequency of heavy drinking SOD						Drinking attitudes SOD					
Step 1	Age	.04	.04 ^{NS}	0.11	1.33 ^{NS}		Age	.08	.08**	0.15	1.79 ^{NS}
	Gender			0.17	2.00*		Gender			0.25	3.01*
Step 2	Age	.34	.30***	-0.06	-0.82 ^{NS}		Age	.17	.11***	0.04	0.52 ^{NS}
	Gender			0.04	0.5 ^{NS}		Gender			0.17	2.1*
	7-day consum			-0.59	-7.96***		7-day consum			-0.36	-4.38***
Step 3	Age	.37	.03*	-0.08	-1.12 ^{NS}		Age	.20	.04 [†]	0.04	0.49 ^{NS}
	Gender			0.02	0.29 ^{NS}		Gender			0.14	1.74 ^{NS}
	7-day consum			-0.60	-8.17***		7-day consum			-0.34	-4.22***
	BIDR SD			0.16	2.22*		BIDR SD			0.00	-0.04 ^{NS}
	BIDR IM			0.04	0.48 ^{NS}		BIDR IM			0.19	2.39*

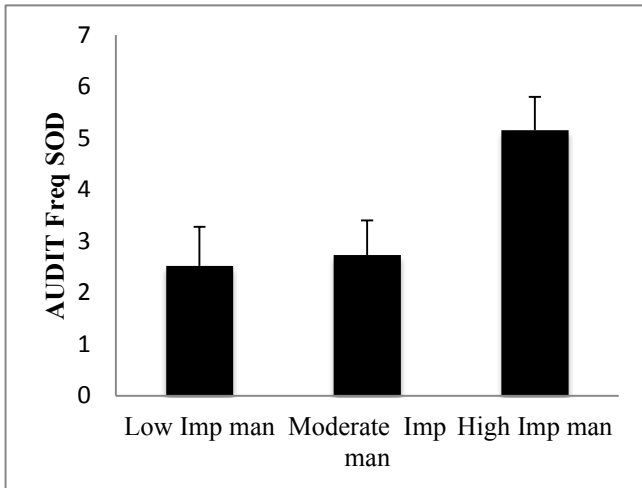


Figure 1a

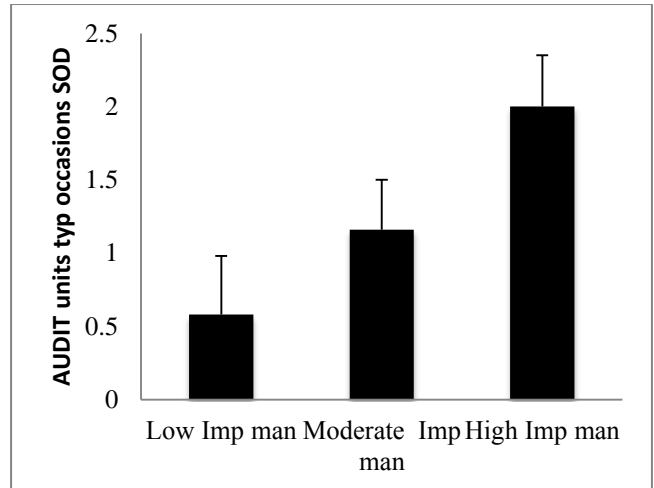


Figure 1b

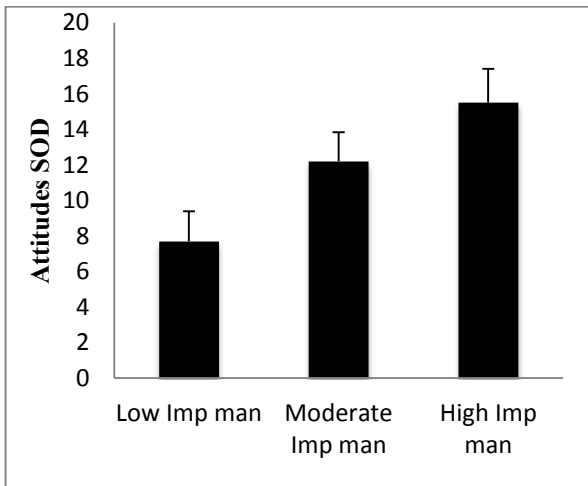


Figure 1c

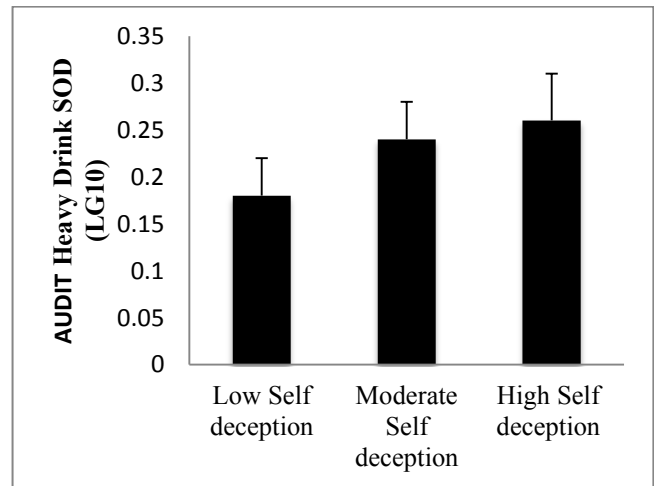


Figure 1d

546

547 Figure 1 (a-d): Mean self-other difference values at low, moderate and high (1a) impression
 548 management for AUDIT frequency of consumption, (1b) impression management for AUDIT
 549 quantity of units consumed on a typical occasion, (1c) impression management for drinking
 550 attitudes index scores, and (1d) self-deception for AUDIT frequency of heavy drinking.

Table (1-3) legends

Table 1:

*** $P < .001$; ^{NS} $P > .05$; d = Cohen's d

MT: Multi-target version of the questionnaire; ST^{PERS}: Single-target 'personal' version of the questionnaire; ST^{AS}: Single-target 'average student' version of the questionnaire.

AUDIT frequency of heavy drinking values were log transformed

Table 2:

MT respondents only $N = 142$

* $P < .05$; ** $P < .01$; *** $P < .001$

BIDR SD/IM = Balanced Inventory of Desirable Responding Self-deception/Impression management subscales

AUD Freq/Quan/Heavy = AUDIT Frequency of consumption/AUDIT Typical quantity of units consumed on a drinking occasion/AUDIT Frequency of heavy drinking

7-day consum = recent seven-day consumption.

Gender: Male = 0, female = 1

AUDIT frequency of heavy drinking and the seven-day consumption values were log transformed

Table 3:

MT respondents only $N = 142$

[†] $P = .051$; * $P < .05$; ** $P < .01$; *** $P < .001$; ^{NS} $P > .05$

SOD = self-other difference; BIDR SD/IM = Balanced Inventory of Desirable Responding Self-deception/Impression management subscales.

7-day consum = recent seven-day consumption.

AUDIT frequency of heavy drinking and seven-day consumption values were log transformed.

Figure 1

551 Figure 1 (a-d):

552 Mean self-other difference values at low, moderate and high (1a) impression management for
553 AUDIT frequency of consumption, (1b) impression management for AUDIT quantity of units
554 consumed on a typical occasion, (1c) impression management for drinking attitudes index
555 scores, and (1d) self-deception for AUDIT frequency of heavy drinking.