

MAIN

# The impact of particular safety behaviours on perceived likeability and authenticity during interpersonal interactions in social anxiety disorder

Grishma Dabas<sup>1</sup>, Karen Rowa<sup>2,3\*</sup>, Irena Milosevic<sup>2,3</sup>, David A. Moscovitch<sup>4</sup> and Randi E. McCabe<sup>2,3</sup>

<sup>1</sup>Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, ON, Canada, <sup>2</sup>Anxiety Treatment and Research Clinic, St Joseph's Healthcare, Hamilton, ON, Canada, <sup>3</sup>Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada and <sup>4</sup>Department of Psychology and Centre for Mental Health Research and Treatment, University of Waterloo, Waterloo, ON, Canada

\*Corresponding author. Email: [krowa@stjoes.ca](mailto:krowa@stjoes.ca)

(Received 3 May 2021; revised 12 September 2022; accepted 14 September 2022; first published online 15 November 2022)

## Abstract

**Background:** Safety behaviours are hypothesized to play a vital role in maintaining social anxiety disorder (SAD), in part by orienting socially anxious individuals to adopt an avoidance-based mindset focused on self-protection and self-concealment. Evidence suggests an association between safety behaviour use and negative social outcomes for individuals with SAD. However, research has largely focused on the broad group of safety behaviours, whereas specific subtypes have received less attention.

**Aim:** The present study aimed to further our understanding of the negative interpersonal consequences of specific types of safety behaviours for individuals with SAD by examining whether active, inhibiting/restricting, or physical symptom management safety behaviour use affects perceived likeability and authenticity during a conversation with a stranger.

**Method:** Individuals with SAD ( $n = 29$ ; mean age 35.5 years) and healthy control (non-SAD) participants ( $n = 40$ ; mean age 18.6 years) engaged in a semi-structured social interaction with trained confederates.

**Results:** Participants with SAD were perceived as significantly less likeable and authentic by the confederates, and rated themselves as significantly less authentic compared with those without SAD. The association between group status and likeability was mediated by the use of inhibiting/restricting safety behaviours and the association between group status and participant-rated authenticity was mediated by the use of both inhibiting/restricting and active safety behaviours, but not physical symptom management strategies.

**Conclusions:** These results contribute to a growing literature suggesting that some, but not all, safety behaviours may play an important role in creating the negative social outcomes that individuals with SAD experience.

**Keywords:** authenticity; likeability; safety behaviours; self-concealment behaviours; social anxiety disorder

## Introduction

Social anxiety disorder (SAD) is an excessive fear of being evaluated negatively in social situations where an individual is exposed to possible scrutiny by others (American Psychiatric Association, 2013). When individuals with SAD encounter anxiety-provoking social situations, they often use safety behaviours in an effort to shield themselves from undesirable social outcomes such as negative evaluation, social rejection, or unbearable feelings of anxiety (Clark and Wells, 1995; McManus *et al.*, 2008; Piccirillo *et al.*, 2016; Salkovskis, 1991).

© The Author(s), 2022. Published by Cambridge University Press on behalf of British Association for Behavioural and Cognitive Psychotherapies. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

The broad group of safety behaviours has been subtyped in several ways. One empirically derived conceptualization divides safety behaviours into active behaviours to enhance performance, such as rehearsing what one will say, inhibiting/restricting behaviours, such as staying quiet to avoid attracting attention, and physical symptom management, such as wearing clothing to hide signs of anxiety (Cuming *et al.*, 2009). Another commonly used subtyping strategy is to separate safety behaviours into avoidance behaviours (any behaviour that decreases one's involvement in a situation) and impression management strategies (any behaviour that controls how one is perceived) (Clark and Wells, 1995). Following from the conceptualization of SAD as fears of revealing negative self-attributes to others (Moscovitch, 2009; Rodebaugh, 2009), another subtype of safety behaviours is self-concealment behaviours, which are strategies used to conceal perceived deficits from others and shield oneself from potential criticism. Self-concealment behaviours are identifiable mainly by their function and therefore have significant content overlap with many of the above noted subtypes of safety behaviours (e.g. speaking very little during a conversation may be both an inhibiting/restricting safety behaviour as well as a self-concealment strategy if the function is to conceal one's fear that they are annoying).

Safety behaviours, broadly, are frequently used by socially anxious individuals, but Clark and Wells (1995) hypothesized that safety behaviours and their subtypes actually contribute to the maintenance of social anxiety symptoms in crucial ways. The use of safety behaviours prevents individuals from being exposed to evidence that disconfirms their fears in social situations. Safety behaviours may increase social anxiety through increased self-monitoring. The use of safety behaviours expends valuable attentional resources, which can result in the individual appearing distant and uninterested, thus exacerbating the likelihood of feared social outcomes (see also Rapee and Heimberg, 1997). Safety behaviours may also reduce feelings of authenticity in social interactions, which are already depleted for individuals with SAD compared with controls (Asher and Aderka, 2021) and are associated with decreased interest in further social interactions (Plasencia *et al.*, 2016).

In line with these negative predicted outcomes, numerous studies suggest that individuals with SAD tend to experience more negative social outcomes than non-anxious individuals (e.g. Alden and Taylor, 2004; Alden and Wallace, 1995; Creed and Funder, 1998; Heerey and Kring, 2007; Meleshko and Alden, 1993; Taylor and Alden, 2011; Voncken *et al.*, 2008). Individuals high in social anxiety are treated more negatively by their peers at school (Blöte *et al.*, 2007), are rated as less 'likeable' (e.g. Creed and Funder, 1998; Heerey and Kring, 2007; Voncken and Dijk, 2013), convey less warmth (Alden and Wallace, 1995; Stopa and Clark, 1993), and exhibit poorer social skills (Beidel *et al.*, 2010; Thompson and Rapee, 2002) during social interactions relative to others low in social anxiety. Socially anxious individuals self-disclose less in relationships and in interactions with strangers (e.g. Montesi *et al.*, 2013; Sparrevojn and Rapee, 2009; Voncken *et al.*, 2010), which may contribute to lowered relationship quality, likeability, and desire for future interaction (Alden and Bieling, 1998; Cuming and Rapee, 2010; Meleshko and Alden, 1993; Voncken *et al.*, 2008).

Several studies have shown that the use of a variety of safety behaviours are related to negative interpersonal outcomes. Rowa *et al.* (2015) demonstrated that individuals with SAD were rated by objective observers as performing worse during a public speaking task than healthy controls and individuals with other anxiety disorders, and that self-reported safety behaviour use mediated these observed group differences. McManus and colleagues (2008) showed that conversation partners viewed their interaction with an individual who engaged in safety behaviours as less enjoyable and perceived the individual as less likeable, whereas elimination of safety behaviours led to better ratings by the interaction partner. Similarly, Stangier and colleagues found that individuals with SAD who engaged in safety behaviours were rated as less warm and outgoing during both conversation and speech tasks (Stangier *et al.*, 2006). In a series of studies, Plasencia and colleagues found that individuals with SAD who used safety behaviours perceived themselves as being less authentic during a conversation task (Plasencia *et al.*, 2011). Conversely, dropping safety behaviours increased socially anxious participants' felt

authenticity and promoted improvements in their desire to interact with a conversation partner (Plasencia *et al.*, 2016).

While much of this research has used the broad concept of safety behaviours, some studies have examined the impact of specific safety behaviour subtypes on social outcomes. For example, avoidance behaviours, but not impression management behaviours, were linked with negative overall impressions of socially anxious individuals (Hirsch *et al.*, 2004) and lower interest in future interactions by conversation partners (Plasencia *et al.*, 2011). Similarly, the use of avoidance safety behaviours, but not impression management behaviours, in a conversation task in a non-clinical sample was related to perceptions of higher anxiety and lower likeability as rated by conversation partners (Gray *et al.*, 2019).

Thus, there is good evidence linking safety behaviours, and specifically avoidance safety behaviours, to specific negative social outcomes in SAD, but minimal research has studied the other conceptualizations of safety behaviours (e.g. active, inhibiting/restricting, and physical symptom management) in predicting social outcomes. Similarly, past studies have used various measures to assess likeability, with an assessment of desire for future interaction being more common. However, such measures may be too intimate to capture perceptions of likeability after a brief, first-time conversation. Therefore, the current study also included a more broadly framed likeability measure that might better capture perceptions of likeability following an initial get-to-know-you task. We also included assessments of both a felt sense of authenticity as well as observed authenticity to ensure that we captured this important interpersonal outcome. In this way, our study aims to replicate and extend previous research findings by providing a more nuanced understanding of the relation between specific safety behaviour use and several negative social outcomes.

The present study examined the impact of safety behaviour subtypes on likeability, partner-rated authenticity, and self-reported authenticity during an interaction task for individuals with SAD as well as individuals with low social anxiety. We hypothesized that individuals with SAD would be perceived by their conversation partners as less likeable, less authentic, and less desirable as future interaction partners than non-SAD participants, and that these group differences would be mediated by self-reported inhibiting/restricting safety behaviour use during the conversation, given that inhibiting/restricting safety behaviours share the most overlap with avoidance safety behaviours. We similarly predicted that individuals with SAD would rate themselves as less authentic than non-SAD participants and that these group differences would be mediated by self-reported inhibiting/restricting safety behaviours. Given the minimal research examining physical symptom management safety behaviours and interpersonal outcomes, no particular hypotheses were made about this subtype.

## Method

### Participants

The SAD group consisted of 29 treatment-seeking individuals with a mean age of 35.5 years ( $SD = 11.3$  years). The majority (58.6%) of the sample was female and White (58.6%). Diagnoses were made using the Diagnostic Assessment Research Tool (DART; McCabe *et al.*, 2017), a semi-structured, modular interview that assesses *DSM-5* diagnoses. The DART was administered by graduate-level clinicians who received thorough training in diagnostic assessment that involved observing at least three interviews conducted by experienced interviewers and then completing at least three interviews while being observed by an experienced interviewer. Diagnoses were reviewed and confirmed during a weekly team meeting led by a psychologist with 20 years of experience in diagnostic assessment. All participants received a principal diagnosis of SAD. Participants had an average of 1.5 additional diagnoses (range of 0 to 4). Common additional diagnoses were generalized anxiety disorder, major depressive disorder and persistent depressive disorder.

**Table 1.** Demographic and baseline characteristics across groups

Measure	Group						Test statistic		
	SAD ( <i>n</i> = 29)			Non-SAD ( <i>n</i> = 40)			$\chi^2$	<i>t</i>	<i>p</i>
	Per cent	<i>M</i>	<i>SD</i>	Per cent	<i>M</i>	<i>SD</i>			
Age		35.55	11.33		18.60	1.65		9.35	<.0001
Sex (% female)	58.6			77.5			2.83		.09
Ethnicity <sup>a</sup>							6.27		.18
Black	3.4			5					
White/European	58.6			50					
Asian	10.3			37.5					
Other	0			7.5					
Education <sup>a</sup>							47.90		<.001
Completed high school	3.4			0					
Some university	3.4			100					
Completed university	24.1								
Some grad school	24.1								
Completed grad school	3.4								
Other	14.8								
Relationship status <sup>a</sup>									
Single	27.6			62.5					
Current relationship	44.8			37.5					
SPIN		46.38	12.79		15.10	6.62			

SPIN, Social Phobia Inventory.

<sup>a</sup>Demographic data could not be obtained for eight participants in the SAD group.

Participants were recruited just prior to the commencement of a group CBT programme for SAD at a tertiary care anxiety out-patient clinic situated within a large Canadian hospital.

The non-SAD group consisted of 40 undergraduate students ( $M = 18.6$  years of age,  $SD = 1.6$  years) from an introductory psychology class at a large Canadian university that is affiliated with the hospital from which the clinical sample was drawn. As shown in Table 1, the majority (77.5%) of non-SAD participants identified as female, and half (50%) identified as White. Control participants were first screened for SAD symptoms using the mini-Social Phobia Inventory (Mini-SPIN; Connor *et al.*, 2001) during recruitment. In addition, the full version of the SPIN was administered prior to the interaction task and anyone with a score over 30 was excused from the rest of the experiment. Scores above 30 have been used as cut-off scores in previous research to signify clinically significant levels of social anxiety symptoms in analogue samples of high socially anxious undergraduate participants (e.g. Moscovitch *et al.*, 2011; Moser *et al.*, 2008). Three participants scored above this threshold in the present study, and thus did not complete the experiment.

Demographic and clinical characteristics of the study sample across groups are presented in Table 1.

### Confederates

We used a total of eight confederates: one for the non-SAD sample and seven for the SAD sample. Confederates and participants were not matched for gender or age. The higher number of confederates for the SAD sample resulted from confederate availability constraints but also allowed for a range of interaction dyads that varied in age and gender match. The non-SAD group and confederate were all undergraduate students at the same university. In the SAD group, confederates ranged in age from 20 to 58 years with a mean age of 31.6 years ( $SD = 17.4$  years). All confederates were female. The internal consistency of confederate-rated

measures was good in both the SAD group (DFI,  $\alpha = .94$ ; RLS,  $\alpha = .95$ ; SEQ-C,  $\alpha = .90$ ) and in the non-SAD group (DFI,  $\alpha = .97$ ; RLS,  $\alpha = .94$ ; SEQ-C,  $\alpha = .89$ ). All confederates were trained in a one-hour session prior to seeing participants. Training included mock conversations with each other while being observed by research staff, and confederates were given further instructions or suggestions to ensure their behaviour was standardized.

### Measures

#### *Social Phobia Inventory (SPIN; Connor et al., 2000)*

The SPIN is a 17-item self-report measure of the severity of current social anxiety symptoms. Items are rated on a 5-point scale (1, not at all; 5, extremely). The SPIN has demonstrated good test–retest reliability, internal consistency, and convergent and divergent validity (Connor *et al.*, 2000). In the current sample, the internal consistency was .91 in the SAD sample and .70 for the non-SAD sample.

#### *State anxiety ratings*

This measure asked the participants to rate how anxious they felt in the present moment, on a 100-point scale (0, not at all; 100, as anxious as possible).

#### *Subtle Avoidance Frequency Examination (SAFE; Cuming et al., 2009)*

The SAFE is a well-validated self-report measure of safety behaviours. Participants rated how frequently they engaged in each of the 32 safety behaviours listed during the conversation they just had on a 5-point scale (1, never; 5, always). The items on the SAFE are divided into three subscales: *inhibiting/restricting behaviours*, *active behaviours* and *physical symptoms management behaviours*. In the current sample, the internal consistency for the total score was .90 in the SAD sample and .91 for the non-SAD sample. The internal consistency was .90 for the inhibiting/restricting behaviours subscale, .87 for the active behaviours subscale, and .80 for the physical symptoms management behaviours subscale.

#### *Participant- and confederate report of authenticity (SEQ; Plasencia et al., 2011)*

The Self-Experience Questionnaire (SEQ) consists of four items of how genuine the participant was during the interaction task on a 7-point scale (1, not at all; 7, very much). One version of the questionnaire was administered to the participant exactly as developed by Plasencia *et al.* (2011; SEQ-P). Examples of items on this measure include: ‘*I felt genuine during the interaction task*’ and ‘*I felt I was artificial*’. Another version was adapted for this study and administered to the confederate, who provided a rating on how genuine they thought the participant was during their conversation (SEQ-C). Items on this version were adjusted to ‘*I felt he/she was genuine during the interaction task*’ and ‘*I felt he/she was artificial*’. The SEQ-P has demonstrated good construct validity and adequate internal consistency ( $\alpha = .76$ ) (Plasencia *et al.*, 2011). In the current sample, the internal consistency was .86 for the participant-reported scale and .91 for the confederate-rated version.

#### *Desire for Future Interaction Scale (DFI; Coyne, 1976)*

The DFI was used to assess the extent to which the confederate hypothetically desired engaging in future social activities with the participant based on the impression they formed during the interaction. The DFI consists of eight items ranging from less to more intimate interactions, and each item is rated on a 7-point scale (1, not at all; 7, very much). Examples of items include ‘*Would you like to meet this person again?*’ and ‘*Would you ask this person for*

advice?. The DFI has demonstrated good reliability (Voncken and Dijk, 2013), and construct validity (Powers and Zuroff, 1988; Taylor and Amir, 2012). In the current sample, the internal consistency was .95.

#### *Reysen Likability Scale (RLS; Reysen, 2005)*

Confederates rated how likeable the participant was on the RLS. Ratings were provided for 11 qualities on a 7-point scale (1, very strongly disagree; 7, very strongly agree). Examples of the items on this scale are '*this person is approachable*' and '*I would like this person as a roommate*'. Three items on this measure are similar to the DFI and would be considered more intimate in nature ('I would like this person as a co-worker', 'I would like this person as a roommate' and 'I would ask this person for advice'); however, the majority of the items on the RLS are much less intimate, making the RLS a more appropriate indicator of likeability following an initial get-to-know-you task. The RLS has shown adequate internal consistency, convergent validity and divergent validity in prior research (Reysen, 2005). In the current sample, the internal consistency was .95.

#### **Procedure**

Participants in both groups followed the same experimental procedures. Participants first completed an initial set of questionnaires asking about demographic information and symptoms of social anxiety. Afterwards, they were invited into a testing room where they engaged in a semi-structured interaction task with a confederate. The participant and confederate were instructed to sit directly across from one another at a table. Although the role of the confederate was not explicitly stated to participants, they were led to believe that the confederate was another participant in the study. A measure of state anxiety was obtained from the participant after the instructions were read, just before beginning the interaction task. The confederate pretended to fill out this state anxiety rating and circled the same rating across all participants. The experimenter exited the room during the interaction task. Upon completion of the task, the participant completed a second set of questionnaires about use of safety behaviours, and a felt-sense of authenticity. At the same time, the confederate was relocated to a different room and completed a set of questionnaires about how likeable and authentic they found the participant.

#### **'Get-to-know-you' task**

The task consisted of six pre-determined 'get-to-know-you' questions based on the paradigm developed by Aron *et al.* (1997) and adapted for the current study (for details, see 'Confederates' section above). The original task has been used in previous research on social anxiety (Taylor and Amir, 2012). Questions in this task include: '*Tell your partner a bit about yourself*' and '*For what in your life do you feel most grateful?*'. Participants and confederates were told that they would be engaging in an interaction task to get acquainted with one another. They were asked to take turns answering the six questions, with the confederate starting with the first question. Once they had answered the first question, the participant also answered it, and then proceeded to answer the second question. They were instructed to continue answering all six questions in this manner. No time limit was placed on responses.

To increase consistency across participants, confederates were instructed to answer each question using pre-scripted answers. The original task by Aron *et al.* (1997) was adapted such that confederates were instructed to ask the participant a follow-up question based on their answers to the first two questions only. Other than these planned probes, confederates were directed to ask any additional questions when they found something the participant said interesting or they wanted to learn more, as they would in an un-staged conversation. These

measures were used to establish and promote a natural flow of conversation within the structure of the six pre-determined questions.

### Statistical analyses

To examine our first hypothesis about group differences in desire for future interaction, likeability and authenticity, the data were analysed using four independent samples *t*-tests examining: (1) confederate's desire for future interaction, as measured by the DFI; (2) confederate's rating of likeability, as measured by the RLS; (3) confederate's rating of the participant's genuineness, as measured by the SEQ-C; and (4) participant's rating of their own genuineness, as measured by the SEQ-P.

Multiple mediator models were used to examine whether differences in likeability and authenticity ratings across participant groups were mediated by safety behaviour subtypes; specifically, inhibiting/restricting, active and physical symptom management behaviours (as measured by the SAFE subscale scores) during the interaction task. Mediation analyses were only performed for the outcome measures that showed a significant difference in ratings across groups. All mediation models were conducted using the PROCESS macro for SPSS (Hayes, 2017), which implements bias-corrected bootstrapping procedures (Preacher and Hayes, 2008). Specifically, a 95% bias-corrected confidence interval (CI) was generated using 5000 bootstrap samples. For all mediation models, the independent variable (participant group) was dummy coded with non-SAD entered as the reference group.

### Results

All data analyses were conducted using version 25 of IBM SPSS Statistics.

#### Descriptive statistics

Descriptive statistics across participant groups are provided in Table 1. Complete demographic data could not be obtained for eight participants in the SAD group. In cases where the missing data were not substantial, missing items were imputed using mean substitution.

No significant differences were found between groups in terms of gender,  $\chi^2 = 2.83$ ,  $p = .09$ , or ethnicity,  $\chi^2 = 6.27$ ,  $p = .18$ . However, relative to the SAD group, the non-SAD group was younger,  $t_{67} = 9.35$ ,  $p < .0001$ , and less educated,  $\chi^2 = 47.90$ ,  $p < .001$ . Given significant differences in age between groups, mediation analyses were conducted with and without controlling for age. As the results did not differ when age was controlled for, this paper presents the results without controlling for age. We did not hold education constant in analyses as education level is overtly tied to age in this sample given the status of the comparison group as undergraduate students. As expected, the SAD group had higher SPIN scores than the non-SAD group,  $t_{67} = 13.24$ ,  $p < 0.001$ . The bivariate correlations between the variables of interest in the present study are displayed in Table 2.

#### Hypothesis 1: Differences in perceived likeability and authenticity based on group status

Results indicate that confederate ratings for SAD participants were significantly lower than the non-SAD group on the RLS<sup>1</sup>,  $t_{67} = -2.93$ ,  $p < 0.01$  and the SEQ-C,  $t_{67} = -2.57$ ,  $p = 0.01$ .

<sup>1</sup>Given that participants in the non-SAD group were of a similar age and current life stage to their confederate while participants in the SAD group and their confederates were of diverse ages and life stages, we removed three items from the RLS that are heavily biased towards similar age and life stage and conducted an additional *t*-test with these items removed ('I would like this person as a co-worker', 'I would like this person as a roommate' and 'This person is similar to me'). Group differences remained statistically significant,  $t_{67} = -2.44$ ,  $p = .02$ .

**Table 2.** Pearson correlations for safety behaviour use, likeability, authenticity, and desire for future interaction

Variable	1	2	3	4	5	6	7	8
SAFE total	—							
SAFE-1	.92**	—						
SAFE-2	.96**	.82**	—					
SAFE-3	.75**	.53**	.70**	—				
DFI	-.19	-.26*	-.15	-.04	—			
RLS	-.36**	-.42**	-.30*	-.18	.91**	—		
SEQ-C	-.25*	-.31**	-.20	-.10	.64**	.69**	—	
SEQ-P	-.41**	-.48**	-.36**	-.17	.28*	.26*	.33**	—

SAFE, Subtle Avoidance Frequency Scale; SAFE-1, inhibiting/restricting behaviours subscale; SAFE-2, Active behaviours subscale; SAFE-3, Physical symptom management subscale; DFI, Desire for Future Interaction Scale; RLS, Reysen Likeability Scale; SEQ-C, Self-Experience Questionnaire (confederate rating); SEQ-P, Self-Experience Questionnaire (participant rating).

\*\* $p < .01$ , \* $p < .05$ .

Participant ratings on the SEQ-P were significantly lower for the SAD group compared with the non-SAD group,  $t_{67} = -2.22$ ,  $p = 0.03$ . Effect sizes were moderate. Confederate ratings did not differ across groups on DFI scores,  $t_{67} = -1.37$ ,  $p = 0.17$ . These results and the corresponding effect sizes are summarized in Table 3.

### **Hypothesis 2: Safety behaviour subtypes as mediators of the association between group status and perceived likeability or authenticity**

We explored the association between participant group and likeability ratings, with each of the SAFE subscales (inhibiting/restricting, active, and physical symptom management behaviours) entered as a mediating variable. Only the inhibiting/restricting behaviours subscale was a significant mediator. The total effect model indicated that participant group accounted for a significant amount of the variance in predicting likeability,  $R^2 = .11$ ,  $F_{1,67} = 8.57$ ,  $p < .01$ ; see Fig. 1. First, there was a significant total effect of participant group on likeability scores,  $c$ ;  $B = -7.51$ ,  $SE = 2.6$ ,  $t = -2.93$ ,  $p < .01$ . Second, there was a significant association between participant group and inhibiting/restricting behaviours,  $a$ ;  $B = 13.70$ ,  $SE = 1.74$ ,  $t = 7.85$ ,  $p < .001$ , indicating that participants with SAD engaged in significantly more inhibiting/restricting behaviours relative to those without SAD. Third, when participant group was held constant, there was a significant association between inhibiting/restricting behaviour use and likeability,  $b$ ;  $B = -.41$ ,  $SE = .17$ ,  $t = -2.36$ ,  $p = .02$ . Finally, when the mediator variable was included in the model, participant group no longer predicted likeability scores,  $c'$ ;  $B = -1.88$ ,  $SE = 3.44$ ,  $t = -.55$ ,  $p = .59$ . The indirect effect of participant group on likeability via inhibiting/restricting behaviours was statistically significant: effect =  $-5.63$ , 95% CI  $[-10.86, -.15]$ .

We also examined the association between participant group and authenticity ratings with each SAFE subscale entered as a mediator. None of the SAFE subscales mediated confederate-rated authenticity. For participant-rated authenticity, the inhibiting/restricting behaviours subscale was a significant mediator. The total effect model indicated that participant group accounted for a significant amount of the variance in predicting participant rated authenticity,  $R^2 = .07$ ,  $F_{1,67} = 4.91$ ,  $p < .03$ ; see Fig. 2. First, there was a significant total effect of participant group on participant authenticity ratings,  $c$ ;  $B = -2.24$ ,  $SE = 1.0$ ,  $t = -2.21$ ,  $p = .03$ . Second, there was a significant association between participant group and inhibiting/restricting behaviours,  $a$ ;  $B = 13.70$ ,  $SE = 1.74$ ,  $t = 7.85$ ,  $p < .001$ , indicating that participants with SAD engaged in significantly more inhibiting/restricting behaviours relative to those without SAD. Third, when participant group was held constant, there was a significant association between inhibiting/restricting behaviour use and participant rated authenticity,  $b$ ;  $B = -.25$ ,  $SE = .06$ ,  $t = -3.87$ ,

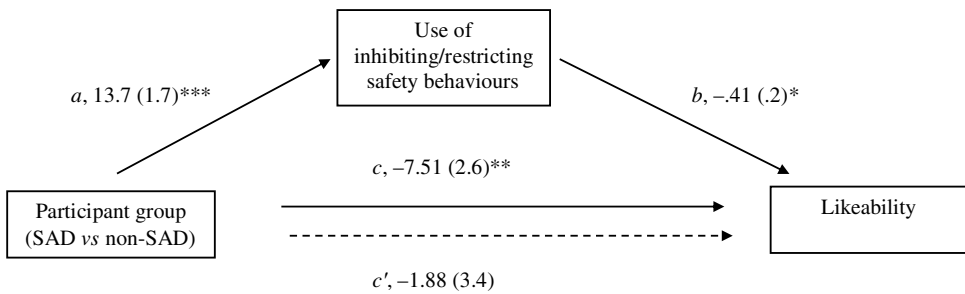


**Table 3.** Group differences in safety behaviour use, likeability, authenticity, and desire for future interaction

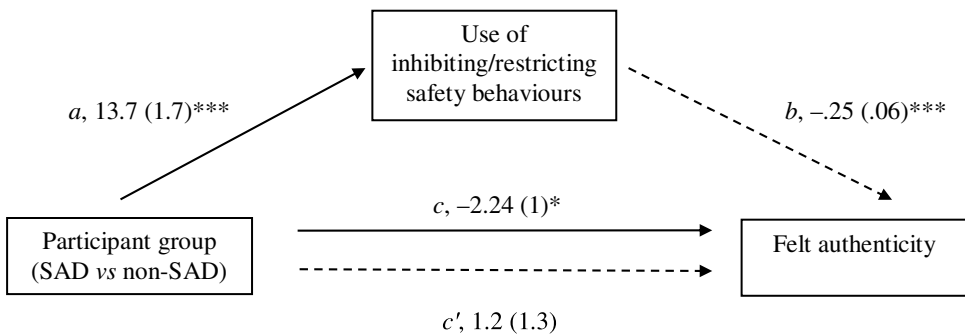
Measure	Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	d.f.	<i>p</i>	<i>d</i>
SAFE total	SAD	29	81.22	19.47	6.81	67	<.0001	1.66
	Non-SAD	40	52.50	15.55				
Inhibiting/restricting behaviours <sup>a</sup>	SAD	29	33.10	8.37	7.85	67	<.0001	1.91
	Non-SAD	40	19.40	6.14				
Active behaviours <sup>a</sup>	SAD	29	36.57	9.23	5.25	67	<.0001	1.28
	Non-SAD	40	25.25	8.55				
Physical symptom management behaviours <sup>a</sup>	SAD	29	11.55	5.19	3.83	67	<.0001	.92
	Non-SAD	40	7.85	2.77				
DFI	SAD	29	35.59	9.19	-1.37	67	.17	0.34
	Non-SAD	40	38.75	9.62				
RLS	SAD	29	49.58	12.16	-2.93	67	.005	0.69
	Non-SAD	40	57.10	9.16				
SEQ-C	SAD	29	24.83	3.89	-2.57	67	.012	0.59
	Non-SAD	40	26.65	1.90				
SEQ-P	SAD	29	22.14	4.85	-2.22	67	.03	0.53
	Non-SAD	40	24.38	3.54				

SAFE, Subtle Avoidance Frequency Scale; DFI, Desire for Future Interaction Scale; RLS, Reysen Likeability Scale; SEQ-C, Self-Experience Questionnaire (confederate rating); SEQ-P, Self-Experience Questionnaire (participant rating).

<sup>a</sup>Measured by SAFE subscales.



**Figure 1.** Mediation model of participant group (dummy coded with non-SAD group as reference) on RLS scores, with SAFE inhibiting/restricting subscale scores as mediators. Unstandardized regression coefficients and standard errors (in parentheses) are provided. Significant effects are represented with continuous lines and non-significant effects are represented with dotted lines. \*\*\**p*<.001, \*\**p*<.01, \**p*<.05.



**Figure 2.** Mediation model of participant group (dummy coded with non-SAD group as reference) on SEQ-P scores, with SAFE inhibiting/restricting subscale scores as mediators. Unstandardized regression coefficients and standard errors (in parentheses) are provided. Significant effects are represented with continuous lines and non-significant effects are represented with dotted lines. \*\*\**p*<.001, \**p*<.05.

$p < .001$ . Finally, when the mediator variable was included in the model, participant group no longer predicted participant rated authenticity,  $c'$ ;  $B = 1.18$ ,  $SE = 1.27$ ,  $t = .92$ ,  $p = .36$ . The indirect effect of participant group on participant rated authenticity via inhibiting/restricting behaviours was statistically significant, effect =  $-3.41$ , 95% CI [ $-5.40$ ,  $-1.46$ ]. Furthermore, the active behaviours subscale was also a significant mediator of this relationship. The total effect model indicated that participant group accounted for a significant amount of the variance in predicting participant rated authenticity,  $R^2 = .07$ ,  $F_{1,67} = 4.91$ ,  $p < .03$ ; see Fig. 3 in Supplemental material. First, there was a significant total effect of participant group on participant authenticity ratings,  $c$ ;  $B = -2.24$ ,  $SE = 1.0$ ,  $t = -2.21$ ,  $p = .03$ . Second, there was a significant association between participant group and active behaviours,  $a$ ;  $B = 11.32$ ,  $SE = 2.16$ ,  $t = 5.25$ ,  $p < .001$ , indicating that participants with SAD engaged in significantly more active safety behaviours relative to those without SAD. Third, when participant group was held constant, there was a significant association between active safety behaviour use and participant rated authenticity,  $b$ ;  $B = -.12$ ,  $SE = .06$ ,  $t = -2.24$ ,  $p = .03$ . Finally, when the mediator variable was included in the model, participant group no longer predicted participant rated authenticity,  $c'$ ;  $B = -.83$ ,  $SE = 1.16$ ,  $t = -.71$ ,  $p = .48$ . The indirect effect of participant group on participant rated authenticity via active behaviours was statistically significant, effect =  $-1.41$ , 95% CI [ $-3.01$ ,  $-.27$ ].

## Discussion

Compared with undergraduate students without SAD, those with SAD were rated as less likeable and authentic by their conversation partners, and also perceived themselves to be less authentic in their interaction with the confederate. These results are broadly consistent with prior research that has demonstrated a significant link between SAD and a variety of negative social outcomes, including likeability (e.g. Alden and Wallace, 1995; McManus *et al.*, 2008; Meleshko and Alden, 1993; Voncken and Dijk, 2013), and self-rated authenticity (Asher and Aderka, 2021; Plasencia *et al.*, 2011). Additionally, results of the current study also add to the growing body of literature suggesting that use of certain types of safety behaviours (specifically inhibiting/restricting and active behaviours) in social interactions may represent a mechanism linking SAD with negative social and interpersonal outcomes (e.g. Furukawa *et al.*, 2009; McManus *et al.*, 2008; Plasencia *et al.*, 2011; Plasencia *et al.*, 2016; Rowa *et al.*, 2015; Stangier *et al.*, 2006; Taylor and Alden, 2011).

Contrary to our expectations, however, group differences on social outcomes did not extend to confederate ratings of desire for future interaction. Past research has been equivocal, with some studies finding differences on desire for future interaction between high and low social anxiety groups (Langer and Rodebaugh, 2013), while others have not (Voncken *et al.*, 2008). These differences may relate to the measurement tool used. Langer and Rodebaugh (2013) asked interaction partners to rate the item 'Would you want to be friends with this person?' on a 7-point scale from *not at all* to *very much*. In comparison, the DFI used in the present study and the study by Voncken *et al.* (2008) contains some items that are more intimate in nature, such as 'Would you invite this person to visit you?'. These types of questions may be too intimate for an initial meeting. Thus, it is possible that the DFI is not an ideal measure to capture likeability after a brief and non-intimate interaction. Alternatively, the non-significant findings on the DFI may also be related to the interaction task being semi-structured. Previous research suggests that socially anxious and non-anxious participants differ less in their social performance in structured vs unstructured social situations (Thompson and Rapee, 2002). This explanation cannot account for significant differences between SAD and non-SAD groups on other variables of interest, but is an interesting consideration.

We hypothesized that group differences on measures of perceived likeability and authenticity would be mediated by inhibiting/restricting safety behaviour use during the conversation based on previous work highlighting avoidance safety behaviours (which overlap the most with inhibiting/restricting behaviours) as more problematic than other types of safety behaviours in understanding negative performance indicators in SAD (Gray *et al.*, 2019; Hirsch *et al.*, 2004; Plasencia *et al.*, 2011). This hypothesis was mainly supported. Self-reported use of inhibiting/restricting, but not active, safety behaviours mediated the relationship between group status and perceived likeability as measured by the Reyson Likeability Scale. Additionally, self-reported use of both inhibiting/restricting and active safety behaviours mediated the relationship between group status and felt authenticity in the get-to-know-you task. However, none of the safety behaviour subtypes mediated the relationship between group status and confederate-rated authenticity. Physical symptom management safety behaviours did not mediate any of the relationships between group status and interpersonal outcomes.

Thus, our findings suggest that inhibiting/restricting and active safety behaviours, in particular, may play a role in the negative interpersonal social outcomes experienced by individuals with SAD. These results extend previous work that has found that avoidance safety behaviours are more problematic than impression management behaviours (Gray *et al.*, 2019; Hirsch *et al.*, 2004; Plasencia *et al.*, 2011) by demonstrating that both active and inhibiting/restricting behaviours, not just inhibiting/restricting behaviours, may contribute to negative interpersonal consequences depending on how these negative consequences are measured. These findings are also consistent with prior research suggesting that safety behaviour use impedes felt authenticity (see Plasencia *et al.*, 2016), supporting the value of measuring multiple interpersonal outcomes to capture the broad range of negative outcomes faced by those with SAD. It is likely that socially anxious individuals who engage in inhibiting/restricting behaviours inadvertently come across as aloof and disinterested in interactions, explaining the link between these types of behaviours and lower ratings of likeability. However, individuals who use active safety behaviours may feel like they are 'putting up a front' that causes them to *feel* disingenuous in social situations, even if they are still perceived as likeable. Another consideration is that both inhibiting/restricting and active safety behaviours may function as self-concealment strategies, used in an effort to hide aspects of the self from others' scrutiny and evaluation. Self-concealment, in all its forms, disrupts the ability for people to make meaningful and authentic connections with others (Moscovitch, 2009).

Physical symptom management safety behaviours did not mediate any of the relationships between group status and interpersonal outcomes. This may mean that this subtype of safety behaviour is truly less influential on likeability and authenticity, but it may also reflect limitations both in how this subtype of safety behaviours was measured as well as the ease with which individuals can use physical symptom management strategies in the moment. For example, things like wearing clothing to hide signs of anxiety or using make-up to hide blushing require pre-planning and may not have been used by participants in this study simply for that reason. Future research should adapt measures of safety behaviours used in experimental situations to ensure that they represent safety behaviours that could be employed in that situation.

Results of the present study revealed that individuals with SAD were seen by interaction partners as appearing less authentic relative to those with low social anxiety, but that this effect was not mediated by the use of any safety behaviours. It is possible, therefore, that group differences in partner-rated authenticity is driven by factors other than SAD participants' use of safety behaviours *per se*. For example, observers are more likely to perceive an individual as authentic if they display behaviours that are positive or moral in their own view (Newman *et al.*, 2014); thus, characteristics of the interaction partner may drive these ratings more than behaviours of the person with SAD. The current study is limited in its ability to assess these hypotheses, but these ideas provide directions for future research.

There were several limitations of the current study. Participant groups differed significantly in age and education level. Although we included age as a covariate in analyses, it is possible that group differences in likeability may have been related to differences in perceived similarity between oneself and one's conversation partner (see Voncken *et al.*, 2008). Likewise, shared student status between the confederate and the group of undergraduates without SAD likely contributed to greater perceived similarity, thus potentially influencing likeability ratings as well. We attempted to control for this by ensuring that confederates in the SAD group represented a broad age range, but this is a limitation. Future studies should use better matched control and SAD groups. Another limitation was that the participants and confederates were not always gender matched during the interaction task. All our confederates were female while the participants were mixed gender. Interactions between same-sex and opposite-sex partners may be qualitatively different (e.g. Carli, 1990), and these differences may have impacted confederates' perceptions of participants. However, a recent SAD study using the same interaction task as in our study has shown that some social outcomes do not depend on gender-matching between participants and confederates (Taylor *et al.*, 2017), suggesting that the impact of gender-matching on social outcomes may not be as large as expected. Furthermore, individuals who engage in higher self-disclosure are typically rated as being more likeable, and one's degree of self-disclosure may, at least in part, depend on one's initial degree of liking of the other individual (Collins and Miller, 1994); however, we did not assess how likeable participants found the confederate, which could have affected how much the participant self-disclosed. In turn, this could have affected the confederate's perception of participants' likeability. The use of a single judge (i.e. confederate) to rate participants on the measures of interest is another limitation. Future studies should consider using other independent observers to corroborate findings. Additionally, although we draw parallels between self-concealment behaviours and inhibiting/restricting and active strategies, the scales used to assess safety behaviours in this study measure behaviour rather than the function of the behaviour. To our knowledge, there is currently no measure of safety behaviours that also assesses the purpose of the behaviour. Thus, one avenue for future research would be to develop scales that assess both the safety behaviour and its function. Another limitation of this study is that the SPIN had low reliability for the non-SAD group for reasons unknown to us. The protocol for the present study also was not pre-registered. Lastly, the small sample size may have rendered the mediation analyses under-powered and may explain the lack of significant effects for age as a covariate as well as the mediations involving confederate-rated and self-rated authenticity. Thus, the mediation data reported in the present study should be interpreted with caution, and future research should confirm these findings in larger samples.

Despite its limitations, results of the present study support the importance of addressing interpersonal variables in the treatment of SAD. Taylor and Alden (2011) developed a modified relational cognitive behavioural therapy (CBT) regimen that focuses on interpersonal functioning to encourage social approach behaviours in individuals with SAD. Relational CBT was shown to be effective in reducing SAD symptoms and improving relationship satisfaction, and these improvements were partially mediated by change in safety behaviour use (Alden *et al.*, 2018). These results, in conjunction with the present study, suggest that limiting both inhibiting and active safety behaviours may play a role in achieving positive therapeutic and interpersonal outcomes. Considering this, clinicians should target and monitor inhibiting and active safety behaviours over the course of treatment, perhaps focusing more on reducing inhibiting/restricting behaviours when the goal of the exposure or experiment is connection, focusing on inhibiting/restricting and active behaviours when the goal is a felt sense of authenticity by the client, or, ideally, focusing on reducing both to maximize positive social outcomes from interactions. Indeed, there is growing awareness among clinicians who specialize in working with SAD that patients' attention should be explicitly drawn to the link between their strategic efforts to self-conceal and the difficulties they experience behaving in

ways that are authentic and likeable. Encouraging patients to experiment with reducing their well-intentioned attempts to present well and noticing the resulting positive effects on interpersonally rewarding outcomes has the potential to help them build meaningful relationships with others and reduce chronic feelings of anxiety, loneliness and social isolation.

**Supplementary material.** To view supplementary material for this article, please visit: <https://doi.org/10.1017/S1352465822000492>

**Data availability statement.** The data that support the findings of this study are available on request from the corresponding author, K.R. The data are not publicly available due to privacy considerations of the patients who participated.

**Acknowledgements.** None.

**Author contributions.** **Grishma Dabas:** Conceptualization (equal), Formal analysis (lead), Methodology (equal), Project administration (equal), Writing – original draft (lead), Writing – review & editing (lead); **Karen Rowa:** Conceptualization (lead), Data curation (lead), Formal analysis (supporting), Methodology (lead), Project administration (lead), Supervision (lead), Writing – original draft (supporting), Writing – review & editing (supporting); **Irena Milosevic:** Conceptualization (supporting), Methodology (supporting), Writing – original draft (supporting), Writing – review & editing (supporting); **David Moscovitch:** Formal analysis (supporting), Methodology (supporting), Writing – original draft (supporting), Writing – review & editing (supporting); **Randi McCabe:** Project administration (supporting), Resources (lead), Writing – original draft (supporting), Writing – review & editing (supporting).

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors

**Conflicts of interest.** The authors declare none.

**Ethical standards.** All materials were reviewed and approved by the Hamilton Integrated Research Ethics Board (1744) and the McMaster University Research Ethics Board (2017 36). All participants provided written informed consent. The authors have abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the BABCP and BPS.

## References

- Alden, L. E., & Bieling, P. (1998). Interpersonal consequences of the pursuit of safety. *Behaviour Research and Therapy*, 36, 53–64. [http://doi.org/10.1016/S0005-7967\(97\)00072-7](http://doi.org/10.1016/S0005-7967(97)00072-7)
- Alden, L. E., Buhr, K., Robichaud, M., Trew, J. L., & Plasencia, M. L. (2018). Treatment of social approach processes in adults with social anxiety disorder. *Journal of Consulting and Clinical Psychology*, 86, 505. <http://doi.org/10.1037/ccp0000306>
- Alden, L. E., & Taylor, C. T. (2004). Interpersonal processes in social phobia. *Clinical Psychology Review*, 24, 857–882. <https://doi.org/10.1016/j.cpr.2004.07.006>
- Alden, L. E., & Taylor, C. T. (2011). Relational treatment strategies increase social approach behaviors in patients with generalized social anxiety disorder. *Journal of Anxiety Disorders*, 25, 309–318. <http://doi.org/10.1016/j.janxdis.2010.10.003>
- Alden, L. E., & Wallace, S. T. (1995). Social phobia and social appraisal in successful and unsuccessful social interactions. *Behaviour Research and Therapy*, 33, 497–505. [https://doi.org/10.1016/0005-7967\(94\)00088-2](https://doi.org/10.1016/0005-7967(94)00088-2)
- American Psychiatric Association. (2013). Anxiety disorders. In *Diagnostic and Statistical Manual of Mental Disorders* (5th edn). <https://doi.org/10.1176/appi.books.9780890425596.744053>
- Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: a procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23, 363–377. <https://doi.org/10.1177/0146167297234003>
- Asher, M., & Aderka, I. M. (2021). How real do you feel? Self- and partner-authenticity in social anxiety disorder. *Journal of Abnormal Psychology*, 130, 166–176. doi:<https://doi.org/10.1037/abn0000622>
- Beidel, D. C., Rao, P. A., Scharfstein, L., Wong, N., & Alfano, C. A. (2010). Social skills and social phobia: an investigation of DSM-IV subtypes. *Behaviour Research and Therapy*, 48, 992–1001. doi: [10.1016/j.brat.2010.06.005](https://doi.org/10.1016/j.brat.2010.06.005)
- Blöte, A. W., Kint, M. J. W., & Westenberg, P. M. (2007). Peer behavior toward socially anxious adolescents: classroom observations. *Behaviour Research and Therapy*, 45, 2773–2779. <https://doi.org/10.1016/j.brat.2007.06.011>
- Carli, L. L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology*, 59(5), 941–951. <https://doi.org/10.1037/0022-3514.59.5.941>
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (eds), *Social Phobia: Diagnosis, Assessment, and Treatment* (pp. 69–93). Guilford Press.
- Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: a meta-analytic review. *Psychological Bulletin*, 116, 457–75. <https://doi.org/10.1037/0033-2909.116.3.457>

- Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Foa, E., & Weisler, R. H. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *British Journal of Psychiatry: Journal of Mental Science*, 176, 379–86. <https://doi.org/10.1192/bjpp.176.4.379>
- Connor, K. M., Kobak, K. A., Churchill, L. E., Katzelnick, D., & Davidson, J. R. T. (2001). Mini-SPIN: a brief screening assessment for generalized social anxiety disorder. *Depression and Anxiety*, 14, 137–140. <https://doi.org/10.1002/da.1055>
- Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology*, 85, 186–193. <https://doi.org/10.1037/0021-843X.85.2.186>
- Creed, A. T., & Funder, D. C. (1998). Social anxiety: from the inside and outside. *Personality and Individual Differences*, 25, 19–33. [https://doi.org/10.1016/S0191-8869\(98\)00037-3](https://doi.org/10.1016/S0191-8869(98)00037-3)
- Cuming, S., & Rapee, R. M. (2010). Social anxiety and self-protective communication style in close relationships. *Behaviour Research and Therapy*, 48, 87–96. <https://doi.org/10.1016/j.brat.2009.09.010>
- Cuming, S., Rapee, R. M., Kemp, N., Abbott, M. J., Peters, L., & Gaston, J. E. (2009). A self-report measure of subtle avoidance and safety behaviors relevant to social anxiety: development and psychometric properties. *Journal of Anxiety Disorders*, 23, 879–883. <https://doi.org/10.1016/j.janxdis.2009.05.002>
- Furukawa, T. A., Chen, J., Watanabe, N., Nakano, Y., Ietsugu, T., Ogawa, S., ... & Noda, Y. (2009). Videotaped experiments to drop safety behaviors and self-focused attention for patients with social anxiety disorder: do they change subjective and objective evaluations of anxiety and performance? *Journal of Behavior Therapy and Experimental Psychiatry*, 40, 202–210. <https://doi.org/10.1016/j.jbtep.2008.08.003>
- Gray, E., Beierl, E. T., & Clark, D. M. (2019). Sub-Types of safety behaviours and their effects on social anxiety disorder. *PLoS ONE*, 14(10), 18. <https://doi.org/10.1371/journal.pone.0223165>
- Gray, E., Beierl, E. T., & Clark, D. M. (2019). Sub-types of safety behaviours and their effects on social anxiety disorder. *PLOS One*, 14, e0223165. <https://doi.org/10.1371/journal.pone.0223165>
- Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (ed T. D. Little) (2nd edn). Guilford Press.
- Heerey, E. A., & Kring, A. M. (2007). Interpersonal consequences of social anxiety. *Journal of Abnormal Psychology*, 116, 125–134. <https://doi.org/10.1037/0021-843X.116.1.125>
- Hirsch, C., Meynen, T., & Clark, D. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory*, 12, 496–506. <https://doi.org/10.1080/09658210444000106>
- Langer, J. K., & Rodebaugh, T. L. (2013). Social anxiety and gaze avoidance: averting gaze but not anxiety. *Cognitive Therapy and Research*, 37, 1110–1120. <https://doi.org/10.1007/s10608-013-9546-z>
- McCabe, R. E., Milosevic, I., Rowa, K., Shnaider, P., Key, B., Pawluk, E., Antony, M. M., & the DART Working Group. (2017). *Diagnostic Assessment Research Tool (DART)*. Hamilton, ON: St Joseph's Healthcare/McMaster University.
- McManus, F., Sacadura, C., & Clark, D. M. (2008). Why social anxiety persists: an experimental investigation of the role of safety behaviours as a maintaining factor. *Journal of Behavior Therapy and Experimental Psychiatry*, 39, 147–161. <https://doi.org/10.1016/j.jbtep.2006.12.002>
- Meleshko, K. G. A., & Alden, L. E. (1993). Anxiety and self-disclosure: toward a motivational model. *Journal of Personality and Social Psychology*, 64, 1000–1009. <https://doi.org/10.1037/0022-3514.64.6.1000>
- Montesi, J. L., Conner, B. T., Gordon, E. A., Fauber, R. L., Kim, K. H., & Heimberg, R. G. (2013). On the relationship among social anxiety, intimacy, sexual communication, and sexual satisfaction in young couples. *Archives of Sexual Behavior*, 42(1), 81–91. <https://doi.org/10.1007/s10508-012-9929-3>
- Moscovitch, D. A. (2009). What is the core fear in social phobia? A new model to facilitate individualized case conceptualization and treatment. *Cognitive and Behavioral Practice*, 16, 123–134. <https://doi.org/10.1016/j.cbpra.2008.04.002>
- Moscovitch, D. A., Gavric, D. L., Merrifield, C., Bielak, T., & Moscovitch, M. (2011). Retrieval properties of negative vs. positive mental images and autobiographical memories in social anxiety: outcomes with a new measure. *Behaviour Research and Therapy*, 49, 505–517. <https://doi.org/10.1016/j.brat.2011.05.009>
- Moser, J. S., Hajcak, G., Huppert, J. D., Foa, E. B., & Simons, R. F. (2008). Interpretation bias in social anxiety as detected by event-related brain potentials. *Emotion*, 8, 693–700. <https://doi.org/10.1037/a0013173>
- Newman, G. E., Bloom, P., & Knobe, J. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, 40, 203–216. <https://doi.org/10.1177/0146167213508791>
- Piccirillo, M. L., Taylor Dryman, M., & Heimberg, R. G. (2016). Safety behaviors in adults with social anxiety: review and future directions. *Behavior Therapy*, 47, 675–687. <https://doi.org/10.1016/j.beth.2015.11.005>
- Plasencia, M. L., Alden, L. E., & Taylor, C. T. (2011). Differential effects of safety behaviour subtypes in social anxiety disorder. *Behaviour Research and Therapy*, 49, 665–675. <https://doi.org/10.1016/j.brat.2011.07.005>
- Plasencia, M. L., Taylor, C. T., & Alden, L. E. (2016). Unmasking one's true self facilitates positive relational outcomes: authenticity promotes social approach processes in social anxiety disorder. *Clinical Psychological Science*, 4, 1002–1014. <https://doi.org/10.1177/2167702615622204>

- Powers, T. A., & Zuroff, D. C.** (1988). Interpersonal consequences of overt self-criticism: A comparison with neutral and self-enhancing presentations of self. *Journal of Personality and Social Psychology*, *54*(6), 1054–1062. <https://doi.org/10.1037/0022-3514.54.6.1054>
- Preacher, K. J., & Hayes, A. F.** (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. <https://doi.org/10.3758/BRM.40.3.879>
- Rapee, R. M., & Heimberg, R. G.** (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, *35*, 741–56. [https://doi.org/10.1016/S0005-7967\(97\)00022-3](https://doi.org/10.1016/S0005-7967(97)00022-3)
- Reysen, S.** (2005). Construction of a new scale: the Reysen Likability Scale. *Social Behavior and Personality*, *33*, 201–208. <https://doi.org/10.2224/sbp.2005.33.2.201>
- Rodebaugh, T. L.** (2009). Hiding the self and social anxiety: the Core Extrusion Schema Measure. *Cognitive Therapy and Research*, *33*, 90.
- Rowa, K., Paulitzki, J. R., Ierullo, M. D., Chiang, B., Antony, M. M., McCabe, R. E., & Moscovitch, D. A.** (2015). A false sense of security: safety behaviors erode objective speech performance in individuals with social anxiety disorder. *Behavior Therapy*, *46*, 304–314. <https://doi.org/10.1016/j.beth.2014.11.004>
- Salkovskis, P. M.** (1991). The importance of behaviour in the maintenance of anxiety and panic: a cognitive account. *Behavioural Psychotherapy*, *19*, 6. <https://doi.org/10.1017/S0141347300011472>
- Sparrevoorn, R. M., & Rapee, R. M.** (2009). Self-disclosure, emotional expression and intimacy within romantic relationships of people with social phobia. *Behaviour Research and Therapy*, *47*, 1074–1078. <https://doi.org/10.1016/j.brat.2009.07.016>
- Stangier, U., Heidenreich, T., & Schermelleh-Engel, K.** (2006). Safety behaviors and social performance in patients with generalized social phobia. *Journal of Cognitive Psychotherapy*, *20*, 17–31. <https://doi.org/10.1891/jcop.20.1.17>
- Stopa, L., & Clark, D. M.** (1993). Cognitive processes in social phobia. *Behaviour Research and Therapy*, *31*, 255–267. doi: 10.1016/0005-7967(93)90024-O
- Taylor, C. T., & Alden, L. E.** (2011). To see ourselves as others see us: an experimental integration of the intra and interpersonal consequences of self-protection in social anxiety disorder. *Journal of Abnormal Psychology*, *120*, 129–141. <https://doi.org/10.1037/a0022127>
- Taylor, C. T., & Amir, N.** (2012). Modifying automatic approach action tendencies in individuals with elevated social anxiety symptoms. *Behaviour Research and Therapy*, *50*, 529–536. doi: 10.1016/j.brat.2012.05.004
- Taylor, C. T., Pearlstein, S. L., & Stein, M. B.** (2017). The affective tie that binds: Examining the contribution of positive emotions and anxiety to relationship formation in social anxiety disorder. *Journal of Anxiety Disorders*, *49*, 21–30. <https://doi.org/10.1016/j.janxdis.2017.03.007>
- Thompson, S., & Rapee, R. M.** (2002). The effect of situational structure on the social performance of socially anxious and non-anxious participants. *Journal of Behavior Therapy and Experimental Psychiatry*, *33*, 91–102.
- Voncken, M. J., Alden, L. E., Bögels, S. M., & Roelofs, J.** (2008). Social rejection in social anxiety disorder: the role of performance deficits, evoked negative emotions and dissimilarity. *British Journal of Clinical Psychology*, *47*, 439–450. <https://doi.org/10.1348/014466508X334745>
- Voncken, M. J., Dijk, C., de Jong, P. J., & Roelofs, J.** (2010). Not self-focused attention but negative beliefs affect poor social performance in social anxiety: an investigation of pathways in the social anxiety-social rejection relationship. *Behaviour Research and Therapy*, *48*, 984–991. <https://doi.org/10.1016/j.brat.2010.06.004>
- Voncken, M. J., & Dijk, K. F. L.** (2013). Socially anxious individuals get a second chance after being disliked at first sight: the role of self-disclosure in the development of likeability in sequential social contact. *Cognitive Therapy and Research*, *37*, 7–17. <https://doi.org/10.1007/s10608-012-9449-4>

---

**Cite this article:** Dabas G, Rowa K, Milosevic I, Moscovitch DA, and McCabe RE (2023). The impact of particular safety behaviours on perceived likeability and authenticity during interpersonal interactions in social anxiety disorder. *Behavioural and Cognitive Psychotherapy* *51*, 46–60. <https://doi.org/10.1017/S1352465822000492>