ALCOHOL RESEARCH UK

Self-Affirmation-Based Enhancement of Risks/Harms-Feedback (SABER-F)

First steps in the development of a novel theory-led internet intervention for heavy drinkers

Sunjeev K Kamboj, Hannah Place, Jess Barton, Stuart Linke, Val Curran and Peter Harris

August 2016

Sunjeev K. Kamboj[1], Hannah Place[1], Jessica A. Barton[1], Stuart Linke[1], Valerie H. Curran[1], Peter R. Harris[2].

1 Research Department of Clinical Educational and Health Psychology, University College London, UK

2 School of Psychology, University of Sussex, UK

Contact: sunjeev.kamboj@ucl.ac.uk

This report was funded by Alcohol Research UK. Alcohol Research UK is an independent charity working to reduce alcohol-related harm through ensuring policy and practice can be developed on the basis of reliable, research-based evidence.
www.alcoholresearchuk.org

CONTENTS

EXECUTIVE SUMMARY

The development of more effective internet-based (eHealth) interventions relies on the application of psychological theory.

Self-affirmation theory proposes that 'self integrity' - the sense of oneself as a moral, consistent and adequate person - is compromised when people receive information that is inconsistent with their behaviour (e.g. providing 'feedback' to a heavy drinker that their consumption is causing them harm).

In this study we examined whether a simple self-affirmation procedure presented in the context of threatening alcohol information (including personalised 'feedback' about cancer risk) could overcome defensiveness in response to this threat, with knock-on effects on intention (to drink less) and drinking behavior.

Although there were no effects on drinking behavior, self-affirmation affected the processing of threatening alcohol-health information in a gender-dependent manner. Weak effects in line with theory and hypotheses were only found in men. Specifically, men showed less defensive and avoidant responding to alcohol-health threat. Women on the other hand, showed paradoxical effects on intention to drink less, which *decreased* in self-affirmed women.

The self-affirmation procedure used here appears to be feasibly delivered via a web-based platform although effects on alcohol-related outcomes may be diluted compared to lab-based procedures. Moreover, an unexpected pattern of gender effects may relate to the population or 'mode of delivery.'

In conclusion, it is feasible to perform self-affirmation procedures in an online environment with at-risk drinkers. However use of internet-based procedures with this population may give rise to (gender-dependent) effects that are substantially diluted compared to lab-based experiments in men.

INTRODUCTION

Cognitive-behavioural cohesion is thought to be maintained by processes that maintain consistency between one's beliefs and behaviour. When important beliefs are challenged, there is a tendency to respond defensively (Chaiken, 1992) as a means of maintaining 'self-integrity' (Steele, 1988). As such, self-worth is retained in the face of psychological threat (e.g. being presented with information that is inconsistent with current beliefs and practises). However, this homeostatic process comes at a cost (Cohen and Sherman, 2014). For example, the tendency to rigidly adhere to dysfunctional beliefs (believing, for example, that advice on smoking or heavy drinking exaggerates the risk or is irrelevant to oneself) is likely to inhibit appropriate self-regulatory behaviour.

Such beliefs survive despite strong contradictory evidence through denial, avoidance or derogation of opposing information (Harris and Napper, 2005). These defensive strategies are a significant barrier to persuasive communication of public health messages. Alcohol consumption is one area of public health in which attempts to modify risk appraisals through threatening messages has been investigated. Overall, such interventions do not appear to modify intention (to drink less) or drinking behaviour (Sheeran et al., 2013). One possibility is that their efficacy is undermined by the defensive processes outlined above.

Although alcohol abuse occurs across the lifespan, high-risk behaviour is especially prevalent in adolescents and college students. Attempts to curb excessive drinking among college students have become a health priority in a number of Western countries (O'Malley and Johnston, 2002). One strategy has been to raise awareness of the risks and costs associated with excessive drinking using brief-feedback-based interventions, which aim to make the negative consequences of excessive drinking more salient (Miller et al., 2013). It is possible that the efficacy of these interventions could be improved by integrating procedures that counteract defensiveness.

Strategies that involve recall of, and reflection upon, personally meaningful values affirm the self as capable, adaptive and moral. These appear to counteract defensive tendencies, allowing greater processing of self-threatening information (Cohen and Sherman, 2014). These effects of 'self-affirmation' have been examined in studies of alcohol-related threatening health information (Armitage et al., 2011, Armitage et al., 2014, Meier et al., 2015, Harris and Napper, 2005, Klein and Harris, 2009, Klein et al., 2011). By and large, these studies tend to support self-affirmation theory by showing, for example, increased fear and intention to reduce consumption in response to information linking alcohol consumption to adverse health effects in self-affirmed participants, although the effects may be moderated upwards or downwards by participants' level of risk (Scott et al., 2013, Harris and Napper, 2005, Klein and Harris, 2009).

Other boundary conditions and mediators may constrain the effectiveness of self-affirmation (Critcher et al., 2010) although these are poorly understood. For example, given that positive pro-social feelings mediate the effects of self-affirmation and that such feelings are proposed to be more easily aroused in women (Crocker et al., 2008), it is possible that gender moderates self-affirmation effects on information-processing. A clarification of gender effects is especially important given the differential risks and harms men and women experience (Nolen-Hoeksema, 2004). Moreover, appraisal of threatening alcohol-related information may differ by gender (Meier et al., 2015).

Furthermore, establishing the effects (or 'effectiveness') of self-affirmation using web-based methods is relevant to the development of eHealth and mHealth interventions, which are

likely to benefit from considering psychological theories of change (Webb et al., 2010, Epton et al., 2013). Moreover, it is likely that continued assessments of self-affirmation procedures will be required to determine suitable parameters for its application. For example opposing effects have been observed depending on the 'strength' of the health message (Klein et al., 2011), potentially giving rise to paradoxical effects on some outcomes (Reed and Aspinwall, 1998).

In the current randomised-controlled experiment we examine the effects of self-affirmation on drinking behaviour and intention to drink less (primary outcomes), and the processing of threatening health-information related to the association between alcohol and cancer (secondary outcomes). The aim was to examine whether the effects typically seen in laboratory-based self-affirmation experiments could be replicated in the context of a less tightly-controlled remote experimental environment. Given the important role of prosocial feelings in mediating self-affirmation effects, and the proposal that this may lead to gender-moderation of effects, the effect of gender was examined across all outcomes.

Method

This randomised controlled, between-subjects experimental study was approved by the University College London Research Ethics Committee.

Participants

Participants (n=292 women; n=236 men) were recruited online specifically using social media sites used by university students from across the UK and randomised to a self-affirmation or control condition. Inclusion criteria were: current UK university student; regular excessive drinking defined as ≥ 4 or ≥ 5 alcohol 'units'/ day (1 unit=8 grams of ethanol) at least four times a week for women and men respectively, consumed alcohol in the past week; age 18-35; fluency in English. Participants were also required to supply a verifiable UK university email address (ending '.ac.uk'), which could only be submitted once during online screening. Each response was required to be linked to a unique IP address. Participants who completed the 1-month follow up were rewarded with a £7 shopping voucher.

Procedure

Eligible participants were emailed a link to the main online experiment. Informed consent was provided online. An online survey programme (Qualitrics, [Provo, Utah, USA]) was used to administer all of the tests and randomise participants to group. Participants were automatically randomised to the self-affirmation (writing about a cherished personal value) or closely-matched control procedure. Blocked randomisation was not used. All participants completed the same sequence of tasks and measures, differing only in the instructions provided for self-affirmation/control tasks.

Demographics, baseline drinking (AUDIT and timeline followback; TLFB) and drinking history were assessed first. Additional exploratory measures followed, typically requiring no more than 2 minutes to complete (e.g. relating to affective response to alcohol images and stages of change), but will not be discussed further here. For the main manipulation we used commonly-employed self-affirmation and control procedures (McQueen and Klein, 2006), in which participants selected one of 11 values which they judged to be the most personally important (self-affirmation) or least personally important but important to another student (control). Immediately after this, participants rated how much love, joyfulness, connectedness and affection they felt (Crocker et al., 2008).

'Generic threat' was then presented, followed by rating of perceived threat; then 'personalised threat' and another threat rating. Ratings of message derogation and avoidance were obtained followed by intention to reduce alcohol-use. At the end of the first online session, participants were given the opportunity to click on a link to a National Health Service (NHS) site containing information about moderation drinking.

Participants were reminded via email to complete the one-week and one-month questionnaires via the online experiment site. Measures of alcohol consumption and intention to reduce consumption were re-assessed at one-week and one-month.

Materials and Measures

Alcohol Use and History

A seven-day timeline-followback (TLFB) procedure (Sobell and Sobell, 1992), was used. An infographic illustrated the content of typical alcohol 'units' (one UK unit=8 g pure alcohol) contained in different beverages (e.g. pint of typical-strength beer). The TLFB is a reliable and valid measure of alcohol consumption, and has additionally been validated via online administration to college students (Pederson et al., 2012). The TLFB was completed before the self-affirmation/control task and at one-week and one-month.

The Alcohol Use Disorders Identification Test (AUDIT) (Babor et al., 2001), a reliable online instrument among young adults (Thomas and McCambridge, 2008) was used to assess levels of harmful drinking. Participants additionally indicated age of first drink ('more than just a sip'), age of regular drinking and family history of 'alcohol-difficulties,' as defined by a list of indices of alcohol-use disorder.

Self-Affirmation and control writing tasks

The self-affirmation task involved writing about one of 11 personally-important values (Sherman et al., 2000). Participants were asked to reflect on and write about how the value had influenced their past behaviour or attitudes and how they use this value in everyday life. With closely matched instructions, participants in the control condition identified the value of least personal importance and wrote about why this value would be important to another student.

The efficacy of the self-affirmation procedure was assessed using ratings of prosocial feelings immediately after the self-affirmation/control procedures (Crocker et al., 2008). Participants rated feelings of 'love,' 'connectedness' and 'affection,' as well as general positive affect ('joy'). Increases in these prosocial feelings are reliably found following self-affirmation procedures (see also Armitage and Rowe, 2011), allowing these to serve as a manipulation check.

Threatening information

'Generic' threatening information consisted of a 211-word outline of the link between alcohol consumption and oral/pharyngeal cancers adapted from information from the UK NHS website on health conditions (http://www.nhs.uk/Conditions). This was followed by a 30 second infomercial graphically depicting the role of alcohol in the development of cancer (www.reducemyrisk.tv). Immediately after this, participants rated the generic (prose plus

infomercial) information on a threat rating scale: 1 (not at all threatening) to 7 (very threatening).

Personalised threatening information was based on participants' gender and weekly alcohol consumption (TLFB), presented as a percentage increase in oral/pharyngeal cancers (Turati et al., 2013). Participants again rated how personally threatening this information was.

Intention to reduce consumption

Participants rated the statement, "I will cut down on the amount of alcohol I drink in the next 7 days:" 1=strongly disagree; 9=strongly agree (Harris and Napper, 2005).

Message derogation, avoidance and acceptance

Message derogation was assessed using items from a previous study (Jessop et al., 2009). Participants first rated the statement "The information about the link between alcohol and cancer was overblown" and then "the message tried to manipulate my feelings". Message avoidance was assessed by asking participants to rate the statement: "my initial reaction was to try and not think about the information" (Jessop et al., 2009). Responses were measured on a scale of 1 to 9 ('strongly disagree' to 'strongly agree').

An indirect measure of message engagement was page dwell-time for the prose, video and personalised threatening information. This is similar to the reading time measure used in previous self-affirmation studies (Reed and Aspinwall, 1998, Klein and Harris, 2009).

Accuracy of responses

At the end of the 1-month follow-up, participants were asked to indicate how 'accurately' they responded to questions across the experiment. To encourage honest responding, the instructions for these questions acknowledged that tiredness and distractions from other tasks may have affected the accuracy of their responses. This was rated on a 0 (not at all accurate) to 100 (very accurate) scale and responses were examined in relation social desirability, as assessed using the short-form (13-item) Marlowe-Crowne scale (Reynolds, 1982).

Data analysis

Data were examined for outliers, defined as values ≥ 3 SD from the mean. Such values were replaced with one plus the largest non-outlying value, except for page-dwell times of ≥100s which were considered spurious and removed. Such adjustments are reflected in the degrees of freedom reported in the statistical analyses. Data was checked for normality to determine appropriateness of parametric statistical analysis.

Between-groups differences in baseline characteristics were assessed using univariate ANOVA, with group and gender as fixed factors. Group and Gender effects on message avoidance, derogation, and threat processing were also examined using univariate ANOVA, again using Group and Gender factors. Repeated measures ANOVA was used to analyse the effects of group and gender on behaviour (Timeline Followback) and intention. Significant interactions were followed up with post-hoc, pairwise Bonferroni corrected tests. Categorical data was analysed using Chi square. The alpha value was set at 0.05, and two-tailed test were used. Appropriate correction for sphericity and inequality of variance was applied and adjusted statistical values (including degrees of freedom) reported.

In common with most internet-based 'intervention' studies, there was a substantial drop off in participation between the first session (n=528), 1 week (n=368; 69.7%) and 1 month follow-up (n=313; 59.3%). On balance, given the amount of missing data and the risk of unreliable replacement of values, list-wise analysis of the existing data was considered preferable to replacement strategies.

All statistical analyses were conducted using SPSS (Version 22) for Windows. Total word count and number of personal pronouns used in the self-affirmation and control tasks were determined using the Linguistic Inquiry and Word Count programme (Pennebaker et al, 2007).

Results

Demographics and alcohol consumption

Characteristics of participants who completed the first experimental session (session-one) are outlined in Table 1. It is worth noting that average AUDIT scores were high (Table 1). As expected, men consumed more alcohol in the preceding week than women [t(526)=5.84, p<0.001]. There were no other significant differences.

Table 1. Demographic and alcohol-relevant variables by gender and group. Except where indicated, values are Mean ± SD

	Control		Self-affirmation	
	Men (n=125)	Women (n=153)	Men (n=111)	Women (n=139)
Age	20.34 (2.87)	20.69 (2.94)	20.68 (3.61)	20.74 (3.24)
AUDIT	17.39 (5.00)	17.48 (5.52)	17.39 (5.25)	17.61 (6.24)
Alcohol (units)	32.24 (17.87)	24.82 (15.18)	35.40 (24.19)	24.32 (14.47)
Age first drink	14.58 (1.94)	14.57 (1.88)	14.41 (1.87)	14.25 (2.38)
Regular drinking*	17.89 (2.04)	17.84 (2.04)	17.83 (1.77)	17.80 (2.57)
Family history**	15.20%	22.22%	12.61%	18.71%

*Age at which starting drinking regularly at current level
** First degree relative with "difficulties with alcohol"

Self-affirmed participants wrote significantly more words (91.50 + 82.77) than non-affirmed participants (74.24 + 53.76) [t(419.50)=2.81, p=0.005, d=0.25]. Moreover self-affirmed participants used a significantly greater number of first person pronouns (15.02 + 8.45) than non-affirmed controls (9.08 + 6.57) [t(468.87)=8.95,p<0.001, d=0.78].

As expected, self-affirmation was associated with higher levels of prosocial/positive feelings (love, connectedness, affection, joy; Table 2) [15]. Since the effect of Group was the same for the four feelings (all were significantly higher in the self-affirmation group, p<0.001), a single composite value is presented. Higher mean prosocial/positive feelings were found in the self-affirmation group (49.80+ 25.22) compared to controls (27.80 + 25.82) [t(525)1=9.88, p<0.001, d=0.86].

Table 2. Prosocial and positive emotions (Mean + SD) assessed after the control/self-affirmation procedure by Gender and Group. Each emotion was rated on a 0-100 scale and values are Mean ± SD.

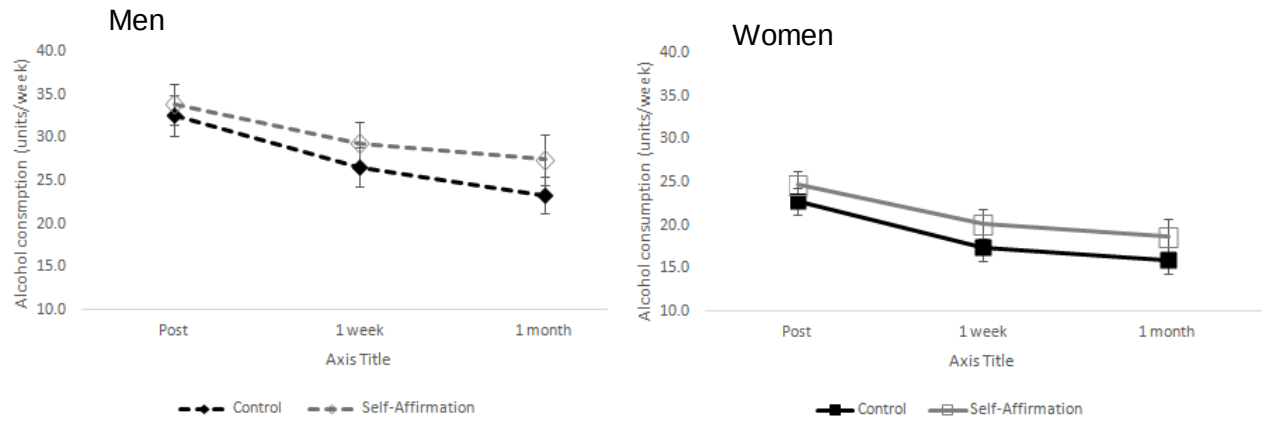
	Control		Self-affirmation	
	Men	Women	Men	Women
Love	26.15 (28.94)	24.75 (28.19)	45.98 (31.31)	52.73 (32.56)
Connection	36.05 (33.48)	31.47 (29.66)	54.68 (30.06)	52.06 (27.87)
Affection	27.42 (28.89)	25.937 (27.41)	47.64 (31.28)	51.78 (30.65)
Joy	29.64 (31.61)	22.46 (25.76)	45.10 (30.04)	47.34 (29.86)

Behavioural outcomes

Alcohol consumption

Based on the TLFB, men consumed significantly more alcohol [$F(1,309)=25.572$, $\eta^2=0.076$, $p<0.0001$], and both groups reduced consumption over time (session-one, one-week, one-month; [$F(2,618)=37.951$, $p<0.0001$, $\eta^2=0.109$]; Figure 1). However, there was no main effect of Group [$F(1,309)=2.181$, $p=0.141$], and no significant interactions involving Gender or Group [F values ≤0.572, p values ≥0.565].

Figure 1. Alcohol Consumption
Alcohol consumption in units per week (Mean + SEM) based on Timeline Followback diary assessment for the previous week's drinking. Men's responses are in the left panel with black diamonds linked by a black dashed line indicating the control group over time, and the grey open diamonds linked by grey dashed line, the self-affirmation group over time. Women's responses for the two groups: control=solid square/solid line; self-affirmation=open grey square/ sold grey line.



Information-seeking

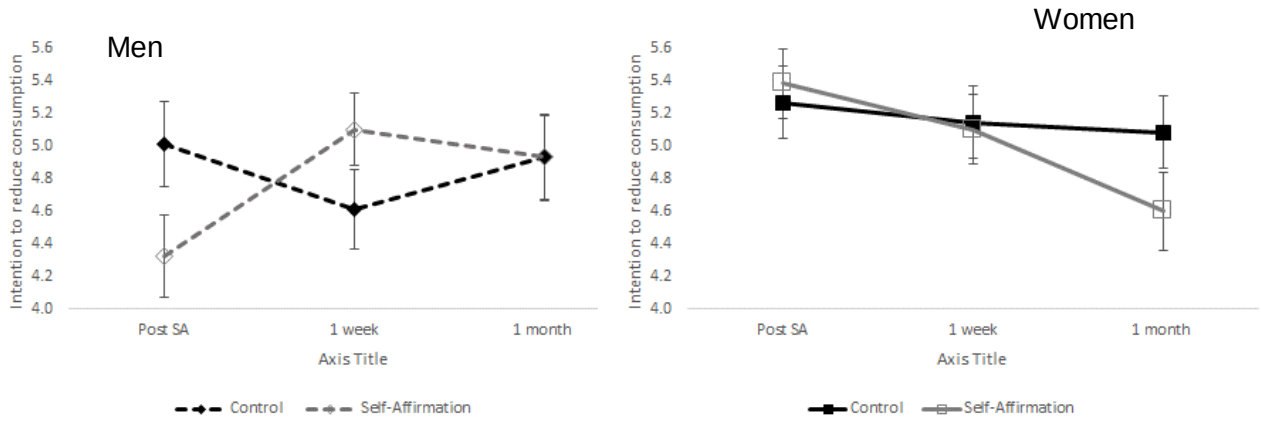
Overall, 21.21% of participants clicked on the link to immediately receive further information about alcohol and health at the end of the experiment. Among women, 26.14 % in control group clicked on the link, compared with 18.71% in the self-affirmation group [$\chi^2(1)=2.304$, $p=0.161$]. Among men, 21.6% in the control group and 26.12% in the self-affirmation group clicked on the link [$\chi^2(1)=0.665$, $p=0.446$].

Intention to reduce alcohol consumption

There were no main effects of Time [F(2,616)=0.63, p=0.531], Group [F(1, 308)=0.27,p=0.606] or Gender [F(1,308)=2.00, p=0.159]. There were however, significant two-way interactions between Time and Group [F(2, 616)=3.19, p=0.042, η_p^2=0.010] and between Time and Gender [F(2, 616)= 5.51, p=0.004, η_p^2=0.018] as well as a three-way Time x Gender x Group interaction [F(2, 616)=5.711, p=0.003, η_p^2=0.018]. Post-hoc Bonferroni-corrected tests compared (i) self-affirmed with control participant at each level of Gender and Time, and (ii) effects at session-one, one-week and one-month at each level of Group and Gender, in a pairwise manner. Somewhat unexpectedly, this showed a trend toward lower intention to reduce consumption in self-affirmed compared to non-affirmed men at session-one (p=0.051). However, considering change in intention across time, men in the self-affirmed group showed an increase in intention to reduce consumption between session-one and one-week (p=0.002), an effect that was sustained (at trend-level) at one-month (compared to session-one; p=0.075; grey dashed line, Figure 2).

Among women, there were no significant differences between self-affirmed and non-affirmed participants at any time-point. However, unlike the increase in intention seen in men, women in the self-affirmation, but not the control group, showed a decrease in intention, at one-month compared to both session-one (p=0.002) and one-week (p=0.039; solid grey line, Figure 2).

Figure 2. Intention to reduce consumption
Ratings of intention to reduce drinking in the next seven days at three time-points ('Post' is immediately after the presentation of self-threatening health information; 1 week and 1 month). Symbols are *Means* ± *SEMs*.

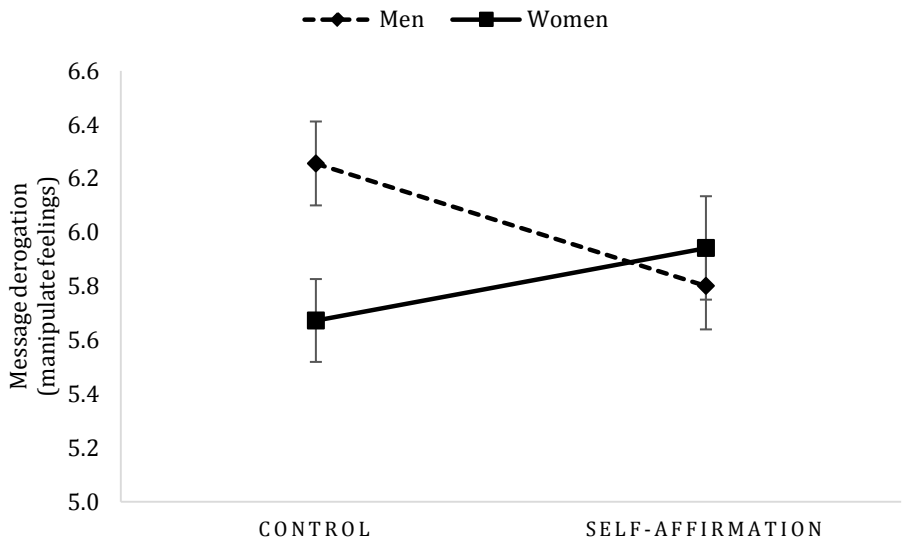


Defensiveness and perceived threat

For the first message-derogation measure ("link between alcohol and cancer was overblown"), there was a main effect of Gender, such that men's scores in the control (4.95 + 1.69) and self-affirmation (4.89 + 1.66) groups were higher than women's (control: 4.41 + 1.71; self-affirmation: 4.60 + 1.64) [F(1,524)=8.09, p=0.005, η_p^2=0.015]). However, there was no effect of Group [F(1,524)=0.18, p=0.670] and no Group x Gender interaction [F(1,524)=0.70, p=0.404].

The second measure of message-derogation ("the message tried to manipulate my feelings") showed no main effects of Gender [F(1,524)=1.78, p=0.183] or Group [F(1, 524)=0.31, p=0.577] but did show a significant Gender x Group interaction [F(1,524)=4.75, p=0.030, η_p^2=0.009] (Figure 3). Post-hoc Bonferroni corrected pair-wise comparisons suggested that underlying this effect was a trend towards lower message derogation in self-affirmed versus non-affirmed men (p=0.065).

Figure 3. Message derogation
Ratings of the message derogation item "the message tried to manipulate my feelings" for the control and self-affirmation groups. The filled diamonds indicate the responses of men (*Mean ± SEM*); solid squares, those of women.



For perceived threat (Table 3), there was a main effect of Threat-type (generic versus personalised), with higher perceived threat in response to personalised information [F(1,477)=15.28, p=0.0001, η_p^2=0.031]. There was a trend-level interaction between Threat-type and Gender [F(1,477)=3.74, p=0.054, η_p^2=0.008], but no main or interaction effect involving Group [F values ≤ 2.95, p values ≥0.1].

Table 3. Perceived threat (Mean + SD) for the two different types of threat: generic and personalised by Gender and Group

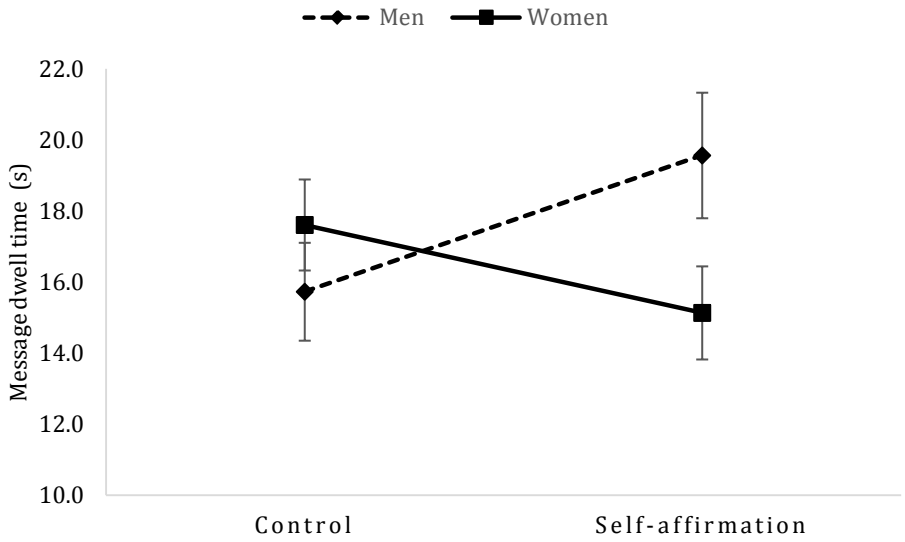
	Control		Self-affirmation	
	Men	Women	Men	Women
Generic	3.93 (1.52)	4.32 (1.53)	4.09 (1.47)	4.38 (137)
Personalised	4.44 (1.80)	4.56 (1.81)	4.37 (1.89)	4.41(1.53)

Message engagement: threat-page dwell-times

Men and women spent similar amounts of time on the prose threat page [F(1, 517)=0.08, p=0.371] and there was no effect of Group [F(1, 516)=0.23, p=0.634]. However, there was Gender x Group interaction [F(1,516)=4.89, p=0.027], which, as depicted in Figure 4, appeared to be driven by a marginal trend for longer dwell times among self-affirmed

compared to non-affirmed men (p=0.071). There appeared to be an opposite pattern in women, although there was no statistical difference between the groups (p=0.196).

Figure 4. Message engagement
Page dwell-time (seconds) for self-threatening prose. Solid diamonds are men's responses (*Mean ± SEM*); solid squares are women's.



For the personalised threatening information, there was a main effect of Gender on page dwell time [F(1, 475)=6.10, p=0.014], such that men viewed this page for longer (mean=10.63 ± 7.64 s) than women (mean=8.97 ± 6.52 s). However, there was no effect of Group [F(1,476)=0.39, p=0.531] and no Group x Gender interaction[F(1,476)=1.03, p=0.310].

The infomercial page was viewed for an average of 36.66 + 20.57. There was no effect of Group [F(1, 506)=0.00, p=0.969] or Gender[F(1,506)=0.43, p=0.512] and no Group x Gender interaction [F(1,506)=0.00, p=0.977].

Estimated accuracy of responses

Participants rated the accuracy of their responses at 80.62 + 15.86 (first session), 79.96 + 14.88 (1 week) and 82. 52 + 15.04 (1 month). These were only weakly correlated with social desirability (Marlow-Crowne scale scores; Mean + SD: 6.10 + 2.69): r values≤0.143 (p values ≥ 0.024). There was no effect of Gender or Group on Marlow-Crowne scores (F values≤0.275, p>0.5). A trend-level interaction between Gender and Group [F(1,524)=2.97, p=0.086), although post-hoc test were non-significant (p values ≥0.132)

DISCUSSION

This study examined the effects of self-affirmation on drinking, intention and health-threat processing in high-risk university students. It builds on previous related work in a number of significant ways. Firstly, we purposively sampled high-risk drinkers. Secondly, we examined moderation by gender, a neglected area in self-affirmation research. Finally, we contributed to the relatively new application of web-based experimentation to self-affirmation research.

Our primary findings were that self-affirmation had no effect on behaviour and that effects on intention were moderated by gender. Secondarily, we found gender-moderated

effects of self-affirmation on threat processing, with suggestive (small, trend-level) effects in men. Although these effects in men are broadly in line with self-affirmation theory, the absence of (or paradoxical) effects in women was not expected.

Our study specifically recruited high-risk drinkers. In line with this goal, previous week drinking-levels were high (~33.5 and ~24.5 units/week for men and women respectively) and substantially higher than most previous studies examining self-affirmation in relation to alcohol-use and related outcomes. On the basis of previous studies, showing that positive self-affirmation effects (on risk perception, intention, affect, message derogation) were only evident among students who drank the equivalent of at least 14 units/week (Scott et al., 2013, Harris and Napper, 2005), similar effects might have been predicted in the current study. However, other studies have found effects in moderate, as opposed to high-risk, student drinkers (Klein and Harris, 2009) or have found significant effects of self-affirmation on threat processing and consumption in low-risk (non-student) drinkers (Armitage et al., 2011). Vastly different participant characteristics between studies therefore contributes to continued uncertainty about the boundary conditions within which self-affirmation is effective. Students are an inherently at-risk group, but to our knowledge, the current study consists of the most severely at-risk sample yet examined in a self-affirmation study. This was determined using a widely-accepted and validated assessment of harmful drinking (the AUDIT) (Babor et al., 2001) as well as a reliable and valid measure of drinking behaviour (Sobell and Sobell, 1992). The AUDIT scores suggest that the sample, on average, consisted of harmful drinkers (i.e. those who are experiencing current harms rather than being at risk of future harms). The current specific findings may therefore reflect our sample characteristics and indicate an upper limit of risk behaviour beyond which self-affirmation becomes less effective (or even counter-productive), at least in women.

The apparently selective (albeit small) effect on message engagement (dwell time on the prose page) and one of the derogation measures in men was not expected. Crocker and colleagues (see also Armitage and Rowe, 2011) showed that prosocial feelings, such as love, explain the relationship between self-affirmation and acceptance of threat (Crocker et al., 2008). They suggested that this relationship may be stronger in women. However, we found no evidence for this: like Crocker et al (2008) we found large differences between groups, but no effect of gender on prosocial feelings.

Effects on intention were complex. Men in the self-affirmation group initially (immediately after the self-affirmation procedure) had lower intention to reduce alcohol compared to the non-affirmed control group. However, self-affirmed men also showed a significant increase in intention to reduce consumption from session-one to one-week. Since there was no pre-task assessment of intention it is impossible to determine whether men in the self-affirmation group had lower baseline intention or suffered an acute paradoxical effect of self-affirmation before recovering at 1 week. Such apparently opposing effects of self-affirmation on intention have been reported previously, although as in the current study, these did not adversely affect behaviour (Reed and Aspinwall, 1998). Temporal effects on intention in women seem clearer, but in the opposite-to-expected direction. In particular, the self-affirmation group seemed to suffer a reduction in intention to reduce alcohol consumption. While this was neither reflected in group differences in intention at any time-point nor drinking behaviour, this finding did seem to conform to a pattern of seemingly opposing effects in women in the current study. These appeared, at least partially, to drive the Gender x Group interactions on message derogation and prose-threat page dwell-times.

Perhaps the most significant difference between previous alcohol-self-affirmation studies and the current study is our use of web-based experimentation. The latter can generate

concerns about participant-engagement and reliability of responses. A number of considerations however, suggest that overall, participants in the current study provided genuine and reliable responses and engaged seriously with the experiment. Firstly, the pattern of responses on ratings of prosocial emotions (Crocker et al., 2008), which was our primary manipulation check were in line with predictions, indicating that the procedure was acutely effective. Secondly, the page dwell-times for the different types of threat information showed the expected pattern of engagement (in terms of sec/page) given the amount of information presented: infomercial>prose>personalised threat information. For example, the dwell-time for the infomercial (>36s), given its duration (30s) suggests that on average, participants viewed it in its entirety before moving to the next page. In addition since university students read up to ~8 words/sec during skim reading (Hewitt and Brett, 2007), the average dwell-time for the prose-threat (17 s), is consistent with at least low-level processing of an average of ~65% of the text. In sum, these findings suggest that self-affirmation procedures can feasibly be tested online, with implications for the design, testing and implementation of web-based interventions for alcohol-use problems among students (Epton et al., 2013).

In fact, our experimental procedure contained certain components that are potentially therapeutically active (e.g. monitoring of use, education about consequences) and our findings suggest that self-affirmation may have some incremental efficacy in certain brief interventions, at least among men. In particular, they may increase engagement in and reduce defensive responding to alcohol-related information or feedback. Improvement in the efficacy of feedback-based intervention is important given that effects of such interventions are typically small, and often difficult to detect in comparison to control groups procedures which also contain therapeutic elements (Bernstein et al., 2010). In fact however, the factors implicated in the widely observed changes (improvements) in drinking-outcomes in control groups (e.g. regression to the mean) may have been particularly evident in a sample of heavy drinkers, such as in the current study.

In sum, our study showed some limited support for self-affirmation theory, but only in men. The lack of effect on most outcomes in women, and more importantly, the apparently paradoxical effects on intention (in women) indicates the need for future alcohol research involving self-affirmation to consider gender moderation. Since the current findings contrast with those of previous studies on self-affirmation - many of which have only examined women - another important research goal is to determine the factors that constrain the effectiveness of self-affirmation (and potentially contribute to paradoxical effects) in women who drink alcohol at harmful levels.

REFERENCES

Armitage, C. J., Rowe, R. (2011). Testing multiple means of self-afffirmation. *British Journal of Psychology* 102, 535-545

Armitage, C. J., Harris, P. R. & Arden, M. A. (2011). Evidence that self-affirmation reduces alcohol consumption: randomized exploratory trial with a new, brief means of self-affirming. *Health Psychology* 30**,** 633-41.

Armitage, C. J., Rowe, R., Arden, M. A. & Harris, P. R. (2014). A brief psychological intervention that reduces adolescent alcohol consumption. *Journal of consulting and clinical psychology* 82**,** 546.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. & Monyeiro, M. G. (2001). The alcohol use disorders identification test. *Guidelines for Use in Primary Care.* Second Edition. World Health Organisation

Bernstein, J. A., Bernstein, E. & Heeren, T. C. (2010). Mechanisms of change in control group drinking in clinical trials of brief alcohol intervention: implications for bias toward the null. *Drug and Alcohol Review* 29**,** 498-507.

Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin* 18**,** 669-679.

Cohen, G. L. & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology* 65**,** 333-371.

Critcher, C. R., Dunning, D. & Armor, D. A. (2010). When self-affirmations reduce defensiveness: timing is key. *Personality and Social Psychology Bulletin* 36**,** 947-59.

Crocker, J., Niiya, Y. & Miashkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive other-directed feelings. *Psychological Scienc,* 19**,** 740-747.

Epton, T., Norman, P., Sheeran, P., Harris, P. R., Webb, T. L., Ciravegna, F., Brennan, A., Meier, P., Julious, S. A. & Naughton, D. (2013). A theory-based online health behavior intervention for new university students: study protocol. *BMC Public Health* 13**,** 107.

Harris, P. R. & Napper, L. (2005). Self-affirmation and the biased processing of threatening health-risk information. *Personlaity and Social Psychology Bulletin* 31**,** 1250-63.

Hewitt, J. & Brett, C. (2007). The relationship between class size and online activity patterns in asynchronous computer conferencing environments. *Computers & Education* 49**,** 1258-1271.

Jessop, D., Simmonds, L. & Sparks, P. (2009). Motivational and behavioural consequences of self-affirmation interventions: A study of sunscreen use among women. *Psychology and Health* 24**,** 529-544.

Klein, W. M. & Harris, P. R. (2009). Self-affirmation enhances attentional bias toward threatening components of a persuasive message. *Psychological Science* 20**,** 1463-1467.

Klein, W. M., Harris, P. R., Ferrer, R. A. & Zajac, L. E. (2011). Feelings of vulnerability in response to threatening messages: Effects of self-affirmation. *Journal of Experimental Social Psychology* 47**,** 1237-1242.

Mcqueen, A. & Klein, W. M. (2006). Experimental manipulations of self-affirmation: A systematic review. *Self and Identity* 5**,** 289-354.

Meier, E., Miller, M. B., Lechner, W. V., Lombardi, N., Claborn, K. R. & Lleffingwell, T. R. (2015). The I Message Among High-Risk College Students. *Journal of American College Health,* 63, 324-39.

Miller, M. B., Leffingwell, T., Claborn, K., Meier, E., Walters, S. & Neighbours, C. (2013). Personalized feedback interventions for college alcohol misuse: An update of Walters & Neighbors (2005). *Psychology of Addictive Behaviors* 27**,** 909.

Nolen-Hoeksema, S. (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical Psychology Review,* 24**,** 981-1010.

O'Malley, P. M. & Johnston, L. D. (2002). Epidemiology of alcohol and other drug use among American college students. *Journal of Studies on Alcohol, Supplement,* 14, 23-39.

Reed, M. B. & Aaspinwall, L. G. (1998). Self-affirmation reduces biased processing of health-risk information. *Motivation and Emotion* 22**,** 99-132.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology* 38**,** 119-125.

Scott, J. L., Brown, A. C., Phair, J. K., Westland, J. N. & Schus, B. (2013). Self-affirmation, intentions and alcohol consumption in students: a randomized exploratory trial. *Alcohol and Alcoholism* 48**,** 458-63.

Sheeran, P., Harris, P. R. & Epton, T. (2013). Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies. *Psychological Bulletin 140, 511-543*

Sherman, D. A., Nelson, L. D. & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin* 26**,** 1046-1058.

Sobell, L. C. & Sobell, M. B. 1992. Timeline follow-back. *Measuring alcohol consumption.* Springer.

Steel, C. M. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology* 21**,** 261-302.

Thomas, B. A. & McCambridge, J. (2008). Comparative psychometric study of a range of hazardous drinking measures administered online in a youth population. *Drug and Alcohol Dependence* 96**,** 121-127.

Turati, F., Garavello, W., Tramacere, I., Pelucchi, C., Galeone, C., Bagnardi, V., Corroa, G., Islami, F., Fedirko, V. & Boffetta, P. (2013). A meta-analysis of alcohol drinking and oral and pharyngeal cancers: results from subgroup analyses. *Alcohol and Alcoholism* 48**,** 107-118.

Webb, T., Joseph, J., Yardley, L. & Michie, S. (2010). Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet research,* 12**,** e4.