Sex, Drugs, and Reckless Driving: Are Measures Biased Toward Identifying Risk-Taking in Men?

Thekla Morgenroth (corresponding author)

T.Morgenroth@exeter.ac.uk

Cordelia Fine

University of Melbourne

Michelle K. Ryan

University of Exeter and University of Groningen

Anna E. Genat

University of Melbourne

Acknowledgements: This work was supported by an internal competitive grant from the Melbourne Business School, University of Melbourne, awarded to Cordelia Fine, who is also grateful for the support of the Women's Leadership Institute Australia. This work was also supported by an internal competitive grant from the College of Life and Environmental Sciences, University of Exeter, awarded to Thekla Morgenroth and Michelle Ryan. Michelle Ryan was also supported by a British Academy Mid-Career Fellowship. The authors would further like to thank Teri Kirby for her helpful comments on an earlier draft of the manuscript.

Abstract

We investigated whether risk-taking measures inadvertently focus on behaviors that are more normative for men, resulting in the overestimation of gender differences. Using a popular measure of risk-taking (DOSPERT) in Study 1 (N=99) we found that conventionally used behaviors were more normative for men, while, overall, newly developed behaviors were not. In Studies 2 (N=114) and 3 (N=124) we demonstrate that differences in normativity are reflected in gender differences in self-reported risk taking, which are dependent on the specific items used. Study 3 further demonstrates that conventional, masculine risk behaviors are perceived as more risky than newly generated, more feminine items, even when risks are matched. We conclude that there is confirmation bias in risk-taking measurement.

Key words: Risk-taking, risk aversion, sex differences, gender differences, gender norms

Risk-taking is often seen as a personality trait that leads to occupational and economic success (Hoffman & Yoeli, 2013). It is also strongly associated with masculinity. Consider, for instance, that business-jargon refers to "big, hairy, audacious goals" (Collins & Porras, 1994), while those too timid to take risks are encouraged to "grow some balls". This association also arises in academic work, where risk-taking is an attribute used to define masculinity (Bem, 1974), and it is argued that risk-taking is a male adaptive trait that increases reproductive success (Herbert, 2015). In line with this association, numerous studies measuring risk-taking behavior, ranging from investment decisions to risky sexual behaviors, drug use, and reckless driving, conclude that men are more willing to take risks than women (e.g., Chen, Baker, Braver & Li, 2000; Hartog, Ferrer-i-Carbonell, & Jonker, 2002; Poppen, 1995; SAMHSA, 2014).

However, there are several reasons to question the universality of gender differences in risk-taking – reasons that leads us to suggest that current measures of risk-taking are biased toward identifying risk-taking in men. First, risk-taking is not a general personality trait. People's willingness to take risks varies across domains, and differences in risk-propensity, including gender differences, can be partly explained by the perceived likelihood of positive and negative consequences (Hanoch, Johnson & Wilke, 2006; Weber, Blais & Betz, 2002). For example, Harris, Jenkins, and Glaser (2006) argue that women's reluctance to engage in risky behaviors, such as gambling or walking home alone at night, can be explained by perceptions of negative outcomes associated with these behaviors and expectations of less enjoyment. Why might this be? Factors such as norms, expectations of success, values, and familiarity affect our choices (e.g., Eccles & Wigfield, 2002; Sunstein, 1996; Wang, Keller & Siegrist, 2011). Importantly, in many risk-taking situations, gender norms and socialization will systematically influence such factors – and, in turn, if and when individuals are willing to take risks.

Take, for example, reckless sexual behavior. Men are often found to be more willing than women to take sexual risks (Poppen, 1995). However, reckless sexual behavior entails different consequences and benefits for men and women. It only holds the risk of unwanted pregnancy for women. Moreover, men are more likely than women to experience pleasure when engaging in casual sex (Armstrong, England, & Fogarty, 2012). Lastly, while a high number of sexual partners is socially acceptable for men, women who behave in the same way risk being labeled "sluts" (Conley, Ziegler, & Moors, 2013).

However, other risks may present a mix of norms and anticipated benefits that are more favorable to women, such as confronting sexist remarks or extreme dieting. Thus, the specific items that researchers choose to measure risk are critical to observed gender differences. A concern is that a *think male – think risk* association will mean that risk-taking that is more typical for women may remain 'under the radar' (Nelson, 2014), biasing the construction of scales that measure risk. This would result in measures that produce data skewed toward identifying risk-taking in men.

We investigated this idea in three studies, drawing on a popular measure of risk-taking, the Domain-Specific Risk-Taking (DOSPERT) scale (Blais & Weber, 2006; Weber et al., 2002). We chose the DOSPERT as it is used to investigate gender differences in risk-taking (Harris et al., 2006; Wilke, Hutchinson, Todd, & Kruger, 2006) and because it includes risky behaviors from different domains, which may reduce its proneness for gender bias given that gender differences vary across domains (Byrnes, Miller, & Schafer, 1999). Thus, a demonstration of bias here would provide strong support for our argument. In Study 1 we test whether behaviors used in the DOSPERT and newly developed risk-taking behaviors are seen as more normative for men or women. In Study 2, we investigate gender differences across these conventional and

new behaviors. Lastly, in Study 3 we examine if these results replicate when conventional and new behaviors are matched in riskiness.

Study 1

In Study 1 we tested whether behaviors conventionally used when measuring risk-taking are seen as more normative for men. We also developed a list of risky behaviors that we anticipated were more normative for women. We predicted that men (vs women) would be rated as more likely to engage in the risky behaviors when using the conventional measure of risk-taking but the opposite would be the case for the new behaviors.

Method

Participants. Participants were 102 Mechanical Turk workers from the US. Three participants were excluded due to failed attention checks, leaving us with 57 women and 42 men. The average age was 35 years (*SD*=12 years). We had 80% power to detect a gender difference effect size of d=.50 (cf. Harris et al., 2006).

Procedure and design. Participants were recruited using Mechanical Turk and were assigned to one of two conditions. They indicated how likely (a) women (compared to men), or (b) men (compared to women), would be to engage in different behaviors. Items were randomized and contained conventional and new items. The study thus had a 2 (Version: Conventional vs. New) X 2 (Target gender: Female vs. Male) mixed design with repeated measures on the first factor.

Measures. The conventional items were based on the DOSPERT (Blais & Weber, 2006). We included risky behaviors from four DOSPERT domains: financial, health and safety, recreational, and social risk-taking (see Harris et al., 2006). We also included new risky behaviors we judged to be more gender-neutral or stereotypically associated with women (see

Table 1). Initial items were generated by two of the authors with the criteria that behaviors should be seen as both more normative for women and as risky, drawing on information regarding gender differences in participation. These suggestions were discussed and revised until all authors agreed.

As a measure of normativity, participants indicated the probability that a man (or woman) would engage in this behavior, on 5-point scales from 1 (much less likely) to 5 (much more likely).[1]

[1] In Studies 1 and 2 we also attempted to replicate Harris' findings concerning the mediating role of negative and positive consequences. We did not replicate the findings, but results can be found in the online supplement.

Table 1

Conventional and New Items Sorted by Domain

Domain	Conventional items (adapted from Blais & Weber, 2006; Harris et al., 2006)	New items
Financial	1. Betting a day's income on the outcome of a sporting event (e.g. baseball, soccer, or football). 2. Betting a day's income at a high stake poker game. 3. Gambling a week's income at a casino.	1. Spending the amount of money she/he would spend on a meal out on scratch cards. 2. Buying a flight from a less reliable airline that often cancels its flights but is 50% cheaper when flying to an important event (which she/he will miss if the flight is canceled). 3. Ordering expensive clothes online when they are on sale, saving 50% but she/he cannot return the clothes if they don't fit
Health and Safety	1. Not wearing a seatbelt as a passenger in the front seat of a car 2. Not wearing a helmet when riding a motorcycle 3. Exposing herself/himself to the sun without using sunscreen 4. Walking home alone at night in a somewhat unsafe area of town	1. Donating a kidney to a family member 2. Undergoing cosmetic surgery if she/he could 3. Sunbathing or using sunbeds on a regular basis 4. Going on an extreme diet to lose weight
Recreational	1. Going whitewater rafting during rapid water flows in the spring 2. Periodically engaging in a dangerous sport (e.g. mountain climbing or sky diving) 3. Piloting her/his own small plane, if she/he could	1. Going horseback riding 2. Playing netball competitively 3. Going rollerblading
Social	1. Admitting that her/his tastes are different from those of her friends 2. Defending an unpopular issue that she/he believes in at a social occasion 3. Arguing with a friend about an issue on which she/he has a very different opinion 4. Interrupting a work meeting or class to ask for clarification on an issue	1. Cook an impressive but difficult dish for a very important dinner party 2. Starting an online petition on a social justice issue 3. Asking her/his boss or supervisor for more flexible working arrangements 4. Confronting a friend or colleague about a sexist remark

Note. Items were presented in random order

Results

Initial analyses revealed that participant gender did not affect ratings of normativity, either by itself or in interaction with target gender. This factor was not included in further analyses. We conducted a 2 (Version: Conventional vs. New) X 2 (Target gender: Female vs. Male) MANOVA with repeated measures on the first factor to test whether men or women were rated as being more likely to engage in risky behaviors when using conventional and new measures of risk-taking in the four domains (see Table 2).

Table 2

Descriptive Statistics for Conventional and New Items

	Conventional items				New items			
	Female target		Male target		Female target		Male target	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Probability of taking financial risks	1.63	0.79	4.51	0.54	2.98	0.56	3.15	0.59
Probability of taking health and safety risks	2.07	0.66	4.27	0.49	4.11	0.60	1.85	0.52
Probability of taking recreational risks	2.07	0.59	4.44	0.48	3.07	0.64	3.03	0.51
Probability of taking social risks	2.83	0.80	3.42	0.68	3.75	0.70	2.17	0.61

Results were in line with our predictions, showing a significant interaction between version and target gender, $F(4, 88) = 146.45$; $p < .001$; $\eta_p^2 = .87$ [.81, .90][2] (see Table 3). We also found a significant effect for target gender, $F(4, 88) = 73.96$, $p < .001$; $\eta_p^2 = .77$ [.67, .82], and a marginal effect for version, $F(4, 88) = 2.26$; $p = .069$; $\eta_p^2 = .09$ [.00, .19]. We followed this analysis with a series of 2 (Version: Conventional vs. New) X 2 (Target gender: Female vs. Male) ANOVAs with repeated measures on the first factor.

For the financial items, we found no main effect for version indicating that engagement in conventional and new behaviors were seen as equally probable. We found a significant main

[2] Values in brackets refer to 95% confidence intervals.

effect for target gender, qualified by an interaction between behavior and target gender. Analyses of simple effects revealed that, in line with predictions, men (vs. women) were rated as more likely to engage in conventional behaviors, $F(1, 95) = 434.19$; $p < .001$, $\eta_p^2 = .82$ [.75, .86], while this gender difference was not present for new behaviors, $F(1, 95) = 2.01$; $p = .159$, $\eta_p^2 = .02$ [.00, .11].

The results from the health and safety domain supported our hypotheses. Neither the main effect for version nor for target gender reached significance, but the interaction was significant. Simple effects analyses showed that for conventional measures of risk-taking, men were rated as more likely to engage in risky behaviors, $F(1, 95) = 342.55$; $p < .001$; $\eta_p^2 = .78$ [.70, .83], but this pattern was reversed for new measures, $F(1, 95) = 386.93$; $p < .001$; $\eta_p^2 = .80$ [.73, .85].

In the recreational domain, we found a main effect for version indicating a lower likelihood of targets engaging in new behaviors, and men were rated more likely to engage in recreational risk-taking than women. These main effects were qualified by a significant interaction between target gender and version. In line with our hypothesis, men were rated as more likely than women to engage in conventional behaviors, $F(1, 96) = 472.06$; $p < .001$; $\eta_p^2 = .83$ [.77, .87], but there was no gender difference for new behaviors, $F(1, 96) = .17$; $p = .684$; $\eta_p^2 < .01$ [.00, .05].

Lastly, in the social domain, we found no main effect for version, but the effect of target gender was significant, indicating that women were rated as more likely to take social risks. The two factors interacted and simple effect analyses revealed that the target gender effect was driven by differences in the new items, where women were rated more likely to take risks, $F(1, 96) =$

142.31; $p < .001$; $\eta_p^2 = .60$ [.47, .68]. The conventional items were again rated as more normative for men, $F(1, 96) = 15.64$; $p < .001$; $\eta_p^2 = .14$ [.04, .27].

Table 3

ANOVA Results for Normativity Ratings

	Version				Gender			Version x Gender		
	df	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Financial	1, 95	0.01	.928	< .01 [.00, .00]	226.39	< .001	.70 [.60, .77]	297.19	< .001	.76 [.67, .81]
Health & safety	1, 95	3.92	.051	.04 [.00, .14]	0.13	.722	< .01 [.00, .05]	539.08	< .001	.85 [.79, .88]
Recreational	1, 96	7.52	.007	.07 [.01, .19]	191.78	< .001	.67 [.56, .74]	254.51	< .001	.73 [.63, .79]
Social	1, 96	3.34	.071	.03 [.00, .13]	19.96	< .001	.17 [.06, .30]	146.31	< .001	.60 [.48, .69]

Note. Values in brackets refer to 95% confidence intervals. Given our sample size and the average correlation between measures (e.g., new and old financial items) we achieved 80% power to find interaction effects of medium size. We achieved >99% power to detect all interaction effects in our study.

Item-by-item analysis. We ran a series of independent sample *t* tests to test whether each of the 14 new items was indeed normative for women or gender-neutral. This was not the case for all items. The items concerned with buying scratch cards, $t(97) = -7.51$, $p < .001$, $d = -1.51$ [−1.94, −1.05], buying a flight from an unreliable airline, $t(97) = -6.75$, $p < .001$, $d = -1.36$ [−1.78, −0.91], and playing netball, $t(96) = -5.88$, $p < .001$, $d = -1.19$ [−1.61, −0.75], were seen as more normative for men than women. However, analyses on effect sizes revealed that these items were still seen as significantly *less* masculine than the conventional items in the respective domains (all $z > 6.32$, $p < .001$). The remaining 11 new items were seen as more normative for women than men with the exception of asking for more flexible working arrangements, which was seen as gender-neutral, $t(97) = 1.35$, $p = .180$, $d = 0.27$ [−0.13, 0.67] (see Online Supplement Material).

Discussion

Our results support the claim that behaviors used in the DOSPERT are more normative for men than women. This was true even for the social domain where women rate themselves as equally likely or more likely to take risks compared to men (Johnson, Wilke, & Weber, 2004; Weber et al., 2002). In contrast, overall, newly developed risk-taking items were seen as more typical of women than men in the health and safety and social domains, and as gender-neutral in the financial and recreational domains. In Study 2 we investigated whether this perceived gender difference in the normativity of behaviors is reflected in self-reported gender differences in risk-taking propensity.

Study 2

Based on Study 1, we predicted that when conventional measures of risk-taking are used, men would rate themselves as more likely to engage in risky behaviors in all four domains, while

when new items are used, women would rate themselves as more likely to engage in risky behaviors in the health and safety and social domains. Further, we expected reduced gender difference for the new items in the financial and recreational domains. We also predicted that the strength of gender normativity would be associated with the magnitude of the gender gap in self-reported risk-taking propensity, such that more masculine risks would be associated with greater male risk-taking.

Method

Participants. Participants were 134 Mechanical Turk workers from the US. We aimed for 50 male and female participants (similar to Study 1) without knowing the gender distribution of future participants. We excluded 9 participants due to failed attention checks and 11 non-heterosexual participants as gender norms differ for members of the LGB community (Niedlich, Steffens, Krause, Settke, & Ebert, 2015). The remaining 52 women and 62 men had an average age of 36 years (*SD*=12 years).

Procedure and design. Participants indicated the likelihood that they would engage in risky behaviors. We presented behaviors from Study 1[3] in random order, adapted to be applicable to the self, and, in one case, edited for clarity (see online supplement). Lastly, participants answered demographic questions. The study had a 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) design with repeated measures on the first factor.

Results

[3] We originally included a few additional risk behaviors in Study 2 (see online supplement). However, as they were not included in Study 1 and we do not know their gender normativity, they were excluded from analyses. The high-stakes poker game item was not included in Study 2.

Initial analyses regarding age revealed that female participants (*M*=40.17; *SD*=13.68) were older than male participants (M=33.39; SD=9.20), $t(112)=3.15$, $p=.002$, and that age was associated with differences in risk-taking. We therefore conducted a 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) MANCOVA controlling for age, with financial, health and safety, recreational, and social risk-taking as the dependent variables, to test whether men were more likely to engage in risky behaviors when using conventional, more masculine measures of risk-taking, while women would be more likely to engage in the new behaviors shown to be more normative for women overall (see Table 4)[4].

In line with predictions, we found that gender and version interacted, $F(4, 102) = 5.34$, $p = .004$, $\eta_p^2 = .17$ [.04, .28]. There was no effect for gender, $F(4, 102) = 1.36$, $p = .254$, $\eta_p^2 = .05$ [.00, .12], and no effect for version $F(4, 102) = 1.84$, $p = .127$, $\eta_p^2 = .07$ [.00, .15]. We then conducted a series of 2 (version: conventional vs. new) × 2 (participant gender: female vs. male) analyses of covariance (ANCOVAs) controlling for age with repeated measures on the first factor to investigate these results in more detail.

[4] Correlations between new and conventional domain-specific risk-taking scores of male and female participants for Study 2 as well as Study 3 can be found in the online supplement. We also provide information about the correlations between the items and overall domain-specific risk-taking.

Table 4

Descriptive Statistics for Likelihood of Taking Risks Depending on Gender and Version

	Conventional (masculine norm)				New			
	Women		Men		Women		Men	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Probability of taking financial risks	1.13	0.51	1.51	0.81	1.94	0.83	1.92	0.78
Probability of taking health and safety risks	1.93	0.86	2.26	0.89	2.09	0.59	1.89	0.64
Probability of taking recreational risks	1.67	0.85	2.14	1.12	2.30	0.98	2.16	1.01
Probability of taking social risks	3.58	0.76	3.32	0.97	3.14	0.85	2.52	0.82

For the financial items, we found no main effect for gender or version (see Table 5). n line with predictions, we found that the two factors interacted and simple effects analyses indicated that while men (vs. women) rated themselves more likely to engage in the conventional masculine behaviors, $F(1, 111) = 5.95$; $p = .016$; $\eta_p^2 = .05$ [.00, .15], this effect of gender was not present in the less masculine new behaviors, $F(1, 111) = 0.03$; $p = .857$; $\eta_p^2 < .01$ [.00, .05], mirroring Study 1.

In the health and safety domain, the likelihood of engaging in behaviors did not differ by version or gender. However, in line with predictions, the factors interacted: men (vs. women) rated themselves marginally more likely to engage in conventional masculine risk-taking behaviors, $F(1, 107) = 3.74$; $p = .056$; $\eta_p^2 = .03$ [.00, .12], while the opposite was true for the new items, $F(1, 107) = 3.63$; $p = .060$; $\eta_p^2 = .03$ [.00, .12].

For recreational behaviors, we found no main effects, but the two factors interacted. Simple effects analyses revealed that, in line with hypothesis, men (vs. women) rated themselves marginally more likely to engage in conventional masculine behaviors, $F(1, 111) = 3.79$; $p = .054$; $\eta_p^2 = .03$ [.00, .12], while there was no difference for new ones, $F(1, 111) = 1.75$; $p = 188$; $\eta_p^2 = .02$ [.00, .09].

Lastly, in the social domain, women rated themselves more likely to take risks, while version in itself did not affect ratings. The factors interacted and simple effects analyses demonstrated that women (vs. men) rated themselves more likely to engage in the new risk-taking behaviors, $F(1, 109) = 12.73$; $p = .001$; $\eta_p^2 = .11$ [.02, .22], mirroring findings from Study 1, but there was no gender difference for the conventional masculine behaviors, $F(1, 109) = 0.89$; $p = .348$; $\eta 2p < .01$ [.00, .07].

We further investigated the effect of norms on gender differences in risk-taking by calculating the correlation between normativity ratings from Study 1 and gender differences in the same behaviors. In line with expectations, we found the two values were highly correlated ($r(27) = .75$, $p < .001$).

Table 5

ANCOVA Results for Likelihood of Taking Risks in Different Domains (Study 2)

		Version			Gender			Version x Gender		
	df	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Financial	1, 111	2.44	.121	.02 [.00, .10]	1.51	.222	.01 [.00, .08]	4.83	.030	.04 [.00, .13]
Health & safety	1, 107	0.26	.615	<.01 [.00, .05]	0.18	.670	<.01 [.00, .04]	10.31	.002	.09 [.01, .20]
Recreational	1, 111	1.44	.233	.01 [.00, .08]	0.15	.700	<.01 [.00, .04]	8.60	.004	.07 [.01, .18]
Social	1, 109	0.91	.343	.01 [.00, .07]	6.71	.011	.06 [.00, .16]	6.18	.014	.05 [.00, .15]

Note. Values in brackets refer to 95% confidence intervals. Analyses control for age. Given our sample size and the average correlation between measures, we achieved 80% power to find small to medium interaction effects. The power we achieved for detecting the four interaction effects were 94% in the financial domain, > 99% in the health and safety domain, > 99% in the recreational domain, and > 99% in the social domain.

Discussion

In line with previous research, we found that when using conventional items to measure risk-taking, men, compared to women, rated themselves as more likely to engage in risky behaviors in the financial domain and marginally more likely to engage in risky behaviors in the health and safety and recreational domains. However, as predicted, these gender differences disappeared or reversed when using behaviors that were gender-neutral or normative for women overall. In the social domain, women (vs men) rated themselves as more likely to take risks, but only when using behaviors normative for women were used to measure risk-taking. Moreover, the extent to which items were normative for men in Study 1 was highly correlated with observed gender differences in engaging in different risky behaviors in Study 2.

However, one could argue that the different findings for conventional and new behaviors were not due to gender norms per se, but due to a confound between objective level of risk and gender-type of risk. Thus, women may be more likely to engage in the new behaviors because they were objectively less risky, rather than because they were more normative for women. To address this, we replicated Study 2 with behaviors matched in levels of riskiness.

Study 3

We aimed to replicate Study 2 using items matched on riskiness. While the probability and impact of a potential costs associated with risky behaviors always differs between people (e.g., based on income or experience), the costs and their probability can be partially quantified for physical and financial risks (e.g., maximum amount that can be lost in a specific gamble) in a way that cannot for social risks (e.g., loss of reputation or social exclusion). We therefore did not include social behaviors and focused on financial and physical risk-taking. We predicted that

men (vs women) would rate themselves more likely to engage in risky behaviors when conventional, more masculine behaviors were used to measure risk-taking, while the opposite would be true for newer, more feminine behaviors. In addition, we tested the hypothesis that activities that are more stereotypically feminine would be perceived as less risky than stereotypically masculine activities, even when riskiness is matched, as masculinity is associated with risk, while femininity is not.

Method

Participants. Participants were 140 Mechanical Turk workers from the US. After excluding non-heterosexual participants and those who failed attention checks, our final sample consisted of 58 women and 62 men with a mean age of 38 years (*SD* = 12 years).

Procedure and design. The procedure was similar to Study 2. In addition to indicating the likelihood that they would engage in risky behaviors, participants rated the riskiness of each behavior, on a 7-point scale from 1 (not at all risky) to 7 (extremely risky) on a separate page. The study had a 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) design with repeated measures on the first factor.

Materials. The behaviors consisted of six items normative for men and six new items[5]. As both the health and safety and recreational domains deal primarily with physical risks, we collapsed them into a single domain called "physical risk" (see Table 6).

Several factors were taken into account when matching items. In the financial domain, we held costs constant (i.e., one day's income), ensured that potential positive outcomes far exceeded costs, and that behaviors came with a risk of not gaining benefits. In the physical

[5] A pilot study (N = 47) confirmed that all conventional items were seen as more normative for men (all $p < .001$) and almost all new items were seen as more normative for women (all $p < .046$). "Going for a 1 ½ hour ride on horseback without wearing a helmet" was seen as more normative for men, but as less masculine than the conventional physical items (see online supplement)

domain, multiple risks could be associated with each behavior (e.g., risk of death, risk of injury), thus we matched items on the risk we judged as the highest concern when engaging in the behaviors. For most behaviors, this was risk of injury, while for skydiving and plastic surgery, we used risk of death. As most sources give risk estimations relative to a specific length of time (e.g., one hour), we adjusted time of engaging in the activity so that the risks for the conventional and new behaviors were equal. For example, riding a motorcycle without a helmet for one hour carries a higher risk of injury than horseback riding without a helmet for the same amount of time (Ball, Ball, Kirkpatrick, & Mulloy, 2006; Fry, Harrison, & Daigneault, 2016; cf Fry et al., 2016; Schulz et al., 2004; Whisman & Hollenhorst, 1999). We therefore adjusted the times to make the behaviors equally risky, adding “for half an hour” to “riding a motorcycle without a helmet” and changed the question about extreme sports such as skydiving to “Taking a skydiving class including one jump”.

Table 6

Risk-Matched Conventional and New Items Sorted by Domain

Domain	Conventional items (adapted from Blais & Weber, 2006; Harris et al., 2006)	New items
Financial	1. Betting a day's income at the horse races	1. Spend a day's income on extremely expensive designer clothes on a disreputable website that offers them for cheap
	2. Betting a day's income at a high-stake poker game	2. Spend a day's income on a spa holiday deal where you can get a full week at a five star hotel, but can't choose the dates and will only be informed last minute, meaning that it is unlikely that you will actually be able to go
	3. Betting a day's income on the outcome of a sporting event.	3. Betting a day's income on the outcome of a dating show such as "The Bachelor"
Physical[a]	1. Riding a motorcycle without a helmet for half an hour	1. Going for a 1 ½ hour ride on horseback without wearing a helmet
	2. Going whitewater rafting at high water in the spring	2. Taking a 4-week cheerleading class
	3. Taking a skydiving class including one jump	3. Getting plastic surgery (knowing that it requires a general anesthetic)

Note. Items were presented in random order. [a]To ensure that risks in this category were purely physical, we told participants to answer these questions "assuming costs didn't matter".

Results

Initial analyses regarding age revealed that female participants (M = 40.40; SD = 11.73) were older than male participants (M = 34.31; SD = 10.11), $t(118) = 3.05$, $p = .003$, and that age was associated with risk-taking. We therefore controlled for age (Table 7).

Table 7

Descriptive Statistics for Likelihood of Engaging in Risky Behaviors and Perceptions of Risk for Men and Women

	Conventional items				New items			
	Women		Men		Women		Men	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Probability of taking financial risks	1.58	1.27	2.16	1.59	1.61	1.01	2.01	1.25
Probability of taking physical risks	2.35	1.48	3.11	1.65	2.64	1.38	2.32	1.30
Perceived riskiness of financial behaviors	5.91	1.29	5.30	1.56	5.20	1.38	4.65	1.34
Perceived riskiness of physical behaviors	5.51	1.08	4.75	1.32	3.94	1.12	3.70	1.11

To test whether the pattern observed in Study 2 would replicate when risk was matched, we conducted a 2 (version: conventional vs. new) × 2 (participant gender: female vs. male) MANCOVA controlling for age with financial and physical risk-taking as the dependent variables. Results indicated that, in line with predictions, there was no overall effect for gender, $F(2, 113) = 1.29$, $p = .280$, $\eta_p^2 = .02$ [.00, .09], or version $F(2, 113) = 2.13$, $p = .123$, $\eta_p^2 = .04$ [.00, .11], and the two factors interacted, $F(2, 113) = 7.76$, $p = .001$, $\eta_p^2 = .12$ [.02, .23].

We then conducted a series of 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) ANCOVAs controlling for age with repeated measures on the first factor. Results regarding the financial items did not replicate our previous findings. Participants rated themselves equally likely to engage in the conventional and new behaviors, there was no effect for gender, and the interaction between the two factors was not significant (see Table 8).

Results regarding physical risks matched our predictions. We found no main effect for gender or version. However, the interaction was significant (see Table 8) and simple effects analyses revealed that, in line with our hypothesis, men (vs. women) rated themselves as marginally more likely to engage in the conventional masculine behaviors, $F(1, 115) = 3.74$; $p =$

.056; η_p^2 = .03 [.00, .12], while the opposite was true for the new behaviors, $F(1, 115) = 2.85$; p = .094; η_p^2 = .02 [.00, .10].

We next investigated whether, despite being matched for risk, the new items would be perceived as less risky. For this, we conducted a 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) MANOVA with repeated measures on the first factor and the financial and physical riskiness as the dependent variables. In line with predictions, the new behaviors were rated as significantly less risky than the conventional ones, $F(2, 116) = 90.94$; $p < .001$; η_p^2 = .61 [.50, .68]. The analysis also revealed a main effect for participant gender indicating that men (vs. women) rated the behaviors as less risky, $F(2, 116) = 5.52$; p = .005; η_p^2 = .09 [.01, .19], and a marginal interaction between version and participant gender, $F(2, 116)$ = 2.99; p = .054; η_p^2 = .05 [.00, .13].

We examined this further in two 2 (Version: Conventional vs. New) x 2 (Participant gender: Female vs. Male) ANOVAs with repeated measures on the first factor. In the financial domain, new items were rated as less risky and this did not depend on gender. Further, men (vs women) rated the riskiness of behaviors in general as lower. In the physical domain, new behaviors were also rated as significantly less risky than the conventional behaviors, and men overall rated behaviors as less risky than women. We also found an interaction, however, simple effects analyses revealed that, in line with predictions, both men and women rated the new behaviors as less risky than the conventional behaviors, $F(1, 118) = 45.57$; $p < .001$; η_p^2 = .28 [.15, .40], and $F(1, 118) = 94.51$; $p < .001$; η_p^2 = .45 [.31, .55], respectively.

Table 8

ANCOVA Results for Probability of Taking Risks and ANOVA Results for Perceived Riskiness

	Version				Gender			Version x Gender		
	df	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Financial (probability)	1, 116	3.09	.081	.03 [.00, .11]	2.65	.106	.02 [.00, .10]	0.12	.730	<.01[.00, .04]
Physical (probability)	1, 115	1.91	.170	.02 [.00, .09]	0.08	.775	<.01 [.00, .03]	16.20	< .001	.12 [.03, .24]
Financial (riskiness)	1, 117	47.24	< .001	.29 [.16, .41]	6.18	.014	.05 [.00, .14]	0.08	.773	<.01 [.00, .03]
Physical (riskiness)	1, 118	136.44	< .001	.54 [.41, .62]	7.55	.007	.06 [.00, .16]	5.27	.024	.04 [.00, .13]

Note. Values in brackets refer to 95% confidence intervals. Analyses control for age. Given our sample size and the correlations between our repeated measures, we achieved 80% power to detect small interaction effects. We only 12% power to find the observed interaction effect in the financial domain. The obtained power for detecting the interaction effect in the physical domain was >99%.

Discussion

In the physical domain, we replicated findings from Study 2 with items matched for riskiness. We found no support for our predictions in the financial domain. In line with predictions, we found that the new behaviors were perceived as less risky than the conventional masculine behaviors, even when risks were matched.

General Discussion

Scientific investigations of differences between women and men have often been charged with bias arising from implicit or explicit assumptions that influence research questions, methods, analysis, and interpretation (Fausto-Sterling, 2008). We provide evidence that researchers overlook more stereotypically feminine forms of risk-taking by inadvertently using more male-typical forms in measurement. We presented three studies demonstrating that a widely used measure of risk-taking, the DOSPERT, is biased toward behaviors more normative for men. When this bias is addressed, gender differences in reported risk-taking disappear, or even reverse, although this pattern was not as consistently observed in Study 3. We showed this both using a diverse range of behaviors (Study 2) as well as a narrow set of items matched on physical or financial riskiness (Study 3).

To be clear, our argument is not that men and women have identical risk-taking profiles: clearly, they do not. Nor is our claim that women are just as likely to take risks as men: current measures, including our own, don't speak to this question. However, our findings support Nelson's (2014) suggestion that more stereotypically feminine forms of risk-taking are overlooked in research, and this has important implications.

First, it has clear implications for interpretation of findings. We have shown that gender differences are contingent on the specific items chosen, providing an additional challenge to

assumptions that risk-taking is a masculine personality trait (see Fine, 2017). Furthermore, our findings demonstrate that conclusions of greater male risk-taking in a particular domain can't be considered to be generalizable to other forms of risk-taking, even *within* that domain of risk.

Moreover, our findings have social implications. It is suggested that gender gaps in occupational representation and success, and wealth, are in part explained by greater male risk-taking (Croson & Gneezy, 2009; Hoffman & Yoeli, 2013), potentially naturalizing such inequalities (see Fine 2017; Nelson, 2014). Thus, the stakes are high when it comes to measurement, and factors such as gender norms and expectations for success, and how these may vary both across and within domains of risk, need to be taken into consideration.

While our research provides valuable insights into the relationship between gender norms, measurement, and gender differences in risk-taking, future research should explore these questions more deeply. It may therefore be beneficial to test these ideas using measurement invariance tests specifically designed to identify bias in psychometric measures. Such tests could reveal the extent to which various current measures of risk-taking are biased toward identifying risk-taking in men – as well as which types of items are particularly prone to this bias - and help to address this bias by developing gender-fair measures of risk-taking. Our research is thus a first step in identifying the need to develop scales that encompass a broader, more gender-balanced range of risk-taking behaviors within domains of interest, a goal that would require more rigorous methods of scale development, based on theoretical considerations and empirical data, than provided here.

Moreover, while our studies relied on self-report, it would be interesting to investigate which behavioral measures of risk are prone to overlooking female risk-taking. For example, researchers have concluded that women are more financially risk-averse than men, using career

choices in stereotypically masculine domains (such as entrepreneurship or finance) as an indicator of risk-taking (Sapienza, Zingales, & Maestripieri, 2009). However, consideration of careers in which women are over-represented, that involve the risk of insufficient success for an adequate income (e.g., modeling, freelancing) might lead to different conclusions.

Future research should also aim to replicate these findings using other measures of risk-taking which may be more normative for men than women, such as the Status-Driven Risk Taking Scale, which focuses on willingness to take physical risks to obtain wealth and success (Ashton, Lee, Pozzobon, & Worth, 2010), and 'harm avoidance' subscales in personality inventories such as the Multidimensional Personality Inventory, which tend to focus on physical risks. As Becker and Eagly (2004) documented, women are as well-represented as men in certain forms of physically risky heroism (living kidney donation, Peace Corps) that are less dependent on physical prowess than acts of heroism in which men dominate (e.g., Carnegie Hero Medal recipients).

Lastly, it should be noted that our results in the financial domain were conflicting. While findings were in line with our predictions in Study 2 this was not the case in Study 3. Here, we found no gender differences at all, regardless of normativity of items. This variation may be potentially due to a sample size which did not provide enough power to detect the particularly small effect size found in Study 3, or due to the overall low ratings of financial risk-taking in our sample. However, it should also be noted that behavioral studies do not uniformly find gender differences in financial risk-taking (Nelson, 2014). Future research should explore this question further.

Conclusion

Risk-taking is strongly associated with men, biasing measures toward identifying risk taking in men. Paying greater attention to female risk-taking is critical for a better understanding of when and why men and women differ in their likelihood to take risks. Moreover, research that challenges, rather than reinforces, cultural assumptions about who takes risks may help to counteract popular conceptions that audacious visions are more likely to come from those who are big and hairy, and that testicles are a necessary condition for courage.

References

Armstrong E. A., England P., Fogarty A. C. (2012). Accounting for women's orgasm and sexual enjoyment in college hookups and relationships. *American Sociological Review, 77*, 435–462.

Ashton M. C., Lee K., Pozzebon J. A., Visser B. A., Worth N. (2010). Status-driven risk taking and the major dimensions of personality. *Journal of Research in Personality, 44*, 734–737.

Ball C. G., Ball J. E., Kirkpatrick A. W., Mulloy R. H. (2007). Equestrian injuries: Incidence, injury patterns, and risk factors for 10 years of major traumatic injuries. *American Journal of Surgery, 193*, 636–640.

Becker S. W., Eagly A. H. (2004). The heroism of women and men. *American Psychologist, 59*, 163–178.

Bem S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162.

Blais A., Weber E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1*, 33–47.

Byrnes J. P., Miller D. C., Schafer W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*, 367–383.

Chen L. H., Baker S. P., Braver E. R., Li G. (2000). Carrying passengers as a risk factor for crashes fatal to 16- and 17-year-old drivers. *Journal of the American Medical Association, 283*, 1578–1582.

Collins J. C., Porras J. (1994). *Built to last: Successful habits of visionary companies*. New York, NY: HarperBusiness

Conley T. D., Ziegler A., Moors A. C. (2013). Backlash from the bedroom: Stigma mediates gender differences in acceptance of casual sex offers. *Psychology of Women Quarterly, 37*, 392–407.

Croson R., Gneezy U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*, 1–27.

Eccles J. S., Wigfield A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132.

Fausto-Sterling A. (2008). *Myths of gender: Biological theories about women and men*. New York, NY: Basic Books.

Fine C. (2017). *Testosterone rex: Myths of sex, science, and society*. New York, NY: Norton.

Fry A. M., Harrison A., Daigneault M. (2016). Micromorts—What is the risk? *British Journal of Oral and Maxillofacial Surgery, 54*, 230–231.

Hanoch Y., Johnson J. G., Wilke A. (2006). Domain specificity in experimental measures and participant recruitment: An application to risk-taking behavior. *Psychological Science, 17,* 300–304.

Harris C. R., Jenkins M., Glaser D. (2006). Gender differences in risk-assessment: Why do women take fewer risks than men? *Judgment and Decision Making, 1,* 48–63.

Hartog J., Ferrer-i-Carbonell A., Jonker N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos, 55*, 3–26.

Herbert J. (2015). *Testosterone: Sex, power, and the will to win*. Oxford: Oxford University Press.

Hoffman M., Yoeli E. (2013). The risks of avoiding a debate on gender differences. *Rady Business Journal, 6,* 6–7.

Johnson J. G., Wilke A., Weber E. U. (2004). Beyond a trait view of risk-taking: A domain-specific scale measuring risk perceptions, expected benefits, and perceived-risk attitude in German-speaking populations. *Polish Psychological Bulletin, 35*, 153–172.

Nelson J. A. (2014). The power of stereotyping and confirmation bias to overwhelm accurate assessment: The case of economics, gender, and risk aversion. *Journal of Economic Methodology, 21*, 211–231

Niedlich C., Steffens M. C., Krause J., Settke E., Ebert I. D. (2015). Ironic effects of sexual minority group membership: Are lesbians less susceptible to invoking negative female stereotypes than heterosexual women? *Archives of Sexual Behavior, 44*, 1439–1447.

Poppen P. J. (1995). Gender and patterns of sexual risk taking in college students. *Sex Roles, 32*, 545–555.

Sapienza P., Zingales L., Maestripieri. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences, 106,* 15268–15273.

Schulz M. R., Marshall S. W., Yang J., Mueller F. O., Weaver N. L., Bowling J. M. (2004). A prospective cohort study of injury incidence and risk factors in North Carolina high school competitive cheerleaders. *The American Journal of Sports Medicine, 32*, 396–405.

Substance Abuse and Mental Health Services Administration (2014). *Results from the 2013 national survey on drug use and health: Summary of national findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Sunstein C. (1996). Social norms and social roles. *Columbia Law Review, 96*, 903–968.

Wang M., Keller C., Siegrist M. (2011). The less you know, the more you are afraid of—A survey on risk perceptions of investment products. *The Journal of Behavioral Finance, 12*, 9–19.

Weber E. U., Blais A., Betz E. N. (2002). A domain specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*, 263–290.

Whisman S. A., Hollenhorst S. J. (1999). Injuries in commercial whitewater rafting. *Clinical Journal of Sports Medicine, 9*, 18–23.

Wilke A., Hutchinson J. M. C., Todd P. M., Kruger D. J. (2006). Is risk taking used as a cue in mate choice? *Evolutionary Psychology, 4,* 367–393.