

ASSESSMENT OF GENDER-RELATED ATTRIBUTES – EMS 1

Table S1

Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's *d*), and Ranks (R) for Sex Differences, Stereotyping, Relevance, and Social Desirability for the masculine Personality Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1

Masculine personality		Sex differences <sup>a</sup>					Stereotyping <sup>b</sup>					Relevance <sup>c</sup>		Social desirability <sup>d</sup>	
Val <sup>1</sup>	Item	Men	Women	Z	<i>d</i>	R	Masculine	Feminine	Z	<i>d</i>	R	Rating	R	Rating	Val <sup>2</sup>
+	Strong	4.64 (1.36)	3.86 (1.43)	-8.52***	-0.56	<b>1</b>	5.77 (1.03)	3.08 (1.40)	-8.04***	-2.19	<b>1</b>	4.36 (2.02)	<b>1</b>	0.86 (1.52)	+
○	<b>Analytical</b>	5.44 (1.29)	4.69 (1.52)	-8.11***	-0.52	<b>2</b>	5.20 (1.59)	3.33 (1.43)	-5.82***	-1.24	<b>16</b>	3.68 (2.11)	<b>14</b>	0.81 (1.59)	+
○	<b>Rational</b>	5.30 (1.35)	4.64 (1.36)	-8.05***	-0.49	<b>3</b>	5.32 (1.41)	3.33 (1.43)	-6.39***	-1.40	<b>10</b>	3.73 (1.98)	<b>13</b>	1.18 (1.64)	+
-	Competitive	4.28 (1.66)	3.49 (1.59)	-7.55***	-0.49	<b>4</b>	5.68 (1.29)	3.77 (1.63)	-6.74***	-1.30	<b>6</b>	3.90 (2.02)	<b>8</b>	0.69 (1.76)	+
+	Bold/Daring	3.97 (1.25)	3.35 (1.26)	-7.54***	-0.49	<b>5</b>	4.95 (1.40)	3.16 (1.29)	-6.48***	-1.33	<b>7</b>	3.60 (1.99)	<b>15</b>	0.27 (1.68)	○
+	Robust	4.93 (1.34)	4.27 (1.42)	-7.16***	-0.47	<b>6</b>	5.41 (1.31)	3.46 (1.36)	-6.95***	-1.46	<b>5</b>	3.79 (2.08)	<b>11</b>	1.39 (1.73)	+
-	<b>Boastful</b>	3.09 (1.47)	2.46 (1.26)	-7.02***	-0.47	<b>7</b>	4.83 (1.64)	2.91 (1.30)	-6.67***	-1.30	<b>8</b>	3.27 (2.11)	<b>18</b>	-1.53 (1.37)	-
○	<b>Willing to take risks</b>	4.13 (1.49)	3.52 (1.49)	-6.16***	-0.41	<b>8</b>	5.27 (1.15)	3.20 (1.26)	-7.21***	-1.72	<b>3</b>	4.00 (1.99)	<b>4</b>	0.13 (1.60)	○
+	Self-confident	4.93 (1.40)	4.50 (1.36)	-5.44***	-0.31	<b>9</b>	5.54 (1.19)	4.46 (1.28)	-5.31***	-0.87	<b>14</b>	3.91 (2.25)	<b>7</b>	1.76 (1.66)	+
+	<b>Courageous</b>	4.67 (1.26)	4.24 (1.26)	-5.31***	-0.34	<b>10</b>	5.47 (1.32)	3.92 (1.21)	-6.22***	-1.22	<b>11</b>	4.09 (2.07)	<b>2</b>	1.17 (1.51)	+
-	<b>Reckless</b>	3.47 (1.46)	3.03 (1.37)	-4.76***	-0.31	<b>11</b>	5.01 (1.10)	2.80 (1.22)	-7.63***	-1.90	<b>2</b>	3.77 (2.06)	<b>12</b>	-0.59 (1.54)	-
-	Ruthless	2.52 (1.36)	2.17 (1.17)	-4.22***	-0.28	<b>12</b>	3.72 (1.66)	2.63 (1.15)	-5.50***	-0.76	<b>20</b>	3.16 (1.94)	<b>19</b>	-1.35 (1.67)	-
+	<b>Pragmatic</b>	4.80 (1.42)	4.51 (1.38)	-3.56***	-0.21	<b>13</b>	5.06 (1.51)	3.39 (1.24)	-6.25***	-1.21	<b>9</b>	3.52 (2.11)	<b>16</b>	0.83 (1.51)	+
-	<b>Controlling</b>	3.64 (1.41)	3.32 (1.45)	-3.56***	-0.22	<b>14</b>	4.69 (1.35)	3.49 (1.44)	-4.75***	-0.86	<b>19</b>	3.97 (2.08)	<b>6</b>	-0.69 (1.76)	-
-	Egotistic	3.55 (1.41)	3.32 (1.36)	-2.68**	-0.17	<b>15</b>	4.49 (1.44)	3.27 (1.43)	-5.50***	-0.85	<b>18</b>	2.95 (2.03)	<b>20</b>	-1.17 (1.65)	-
+	<b>Adventurous</b>	4.80 (1.50)	4.55 (1.55)	-2.33*	-0.16	<b>16</b>	5.37 (1.16)	3.67 (1.26)	-7.09***	-1.40	<b>4</b>	3.87 (2.09)	<b>9</b>	0.33 (1.63)	○
+	Brave	4.76 (1.30)	4.61 (1.31)	-2.05*	-0.12	<b>17</b>	5.45 (1.17)	4.03 (1.44)	-5.98***	-1.08	<b>13</b>	3.86 (1.98)	<b>10</b>	1.29 (1.52)	+
-	<b>Dominant</b>	3.96 (1.41)	3.77 (1.47)	-1.99*	-0.13	<b>18</b>	5.03 (1.27)	3.52 (1.36)	-5.91***	-1.15	<b>12</b>	4.05 (2.11)	<b>3</b>	0.26 (1.58)	○
+	Assertive	4.86 (1.23)	4.72 (1.28)	-1.82 ns	-0.11	<b>19</b>	5.30 (1.15)	4.15 (1.30)	-5.14***	-0.94	<b>17</b>	3.99 (2.16)	<b>5</b>	1.25 (1.67)	+
+	Career-oriented	4.51 (1.63)	4.53 (1.49)	-0.06 ns	0.01	<b>20</b>	5.59 (1.13)	4.25 (1.28)	-5.80***	-1.11	<b>15</b>	3.51 (1.98)	<b>17</b>	1.74 (1.56)	+

Note. Sex difference and stereotyping rank were positively correlated ( $r = .45, p < .05$ ). Relevance rank was not correlated with sex differences ( $r = .05, p = .84$ ) or stereotyping rank ( $r = .30, p = .20$ ); bold items remained in the final version of the questionnaire. Val = Valence; + positive valence; - negative valence; ○ neutral valence; Val<sup>1</sup> = intended valence; Val<sup>2</sup> = rated valence.

ASSESSMENT OF GENDER-RELATED ATTRIBUTES – EMS 1

Table S2

*Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's d), and Ranks (R) for Sex Differences, Stereotyping, Relevance, and Social Desirability for the feminine Personality Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1*

Feminine personality		Sex differences <sup>a</sup>					Stereotyping <sup>b</sup>					Relevance <sup>c</sup>		Social desirability <sup>d</sup>	
Val <sup>1</sup>	Item	Men	Women	Z	d	R	Masculine	Feminine	Z	d	R	Rating	R	Rating	Val <sup>2</sup>
○	Emotional	4.14 (1.49)	5.32 (1.27)	12.14***	0.88	<b>1</b>	3.18 (1.33)	5.79 (1.04)	-7.95***	2.19	<b>1</b>	4.35 (2.01)	<b>5</b>	-0.53 (1.45)	-
-	In need affection	3.65 (1.55)	4.78 (1.43)	11.25***	0.77	<b>2</b>	3.44 (1.49)	5.23 (1.29)	-6.75***	1.28	<b>11</b>	4.12 (2.08)	<b>10</b>	-1.04 (1.40)	-
-	Vulnerable	3.64 (1.48)	4.75 (1.42)	11.23***	0.77	<b>3</b>	3.29 (1.64)	5.23 (1.28)	-6.51***	1.32	<b>10</b>	3.78 (1.90)	<b>16</b>	-1.67 (1.24)	-
+	<b>Compassionate</b>	4.91 (1.42)	5.73 (1.15)	9.87***	0.66	<b>4</b>	3.57 (1.28)	5.80 (1.12)	-7.85***	1.85	<b>3</b>	4.48 (2.01)	<b>2</b>	1.13 (1.66)	+
-	<b>Thin-skinned</b>	3.77 (1.48)	4.68 (1.41)	9.34***	0.66	<b>5</b>	3.33 (1.80)	5.37 (1.12)	-7.15***	1.36	<b>10</b>	3.82 (2.03)	<b>14</b>	-1.44 (1.27)	-
+	Sensitive	4.42 (1.36)	5.21 (1.27)	9.10***	0.61	<b>6</b>	3.45 (1.37)	5.62 (1.08)	-7.60***	1.76	<b>7</b>	4.00 (2.01)	<b>11</b>	0.11 (1.59)	○
+	<b>Delicate</b>	4.63 (1.41)	5.41 (1.23)	9.05***	0.61	<b>7</b>	3.06 (1.25)	5.71 (1.12)	-7.77***	2.23	<b>5</b>	4.36 (2.05)	<b>4</b>	0.55 (1.64)	+
-	<b>Anxious</b>	2.93 (1.40)	3.69 (1.47)	8.02***	0.53	<b>8</b>	2.72 (1.46)	4.39 (1.26)	-6.11***	1.23	<b>19</b>	3.62 (2.16)	<b>18</b>	-1.96 (1.21)	-
+	<b>Loving</b>	4.94 (1.40)	5.59 (1.11)	7.72***	0.54	<b>9</b>	4.31 (1.32)	5.72 (1.16)	-6.41***	1.14	<b>17</b>	4.22 (2.05)	<b>8</b>	0.87 (1.52)	+
+	<b>Caring</b>	4.93 (1.30)	5.52 (1.23)	7.60***	0.47	<b>10</b>	4.45 (1.54)	5.80 (1.00)	-6.41***	1.04	<b>18</b>	4.47 (2.04)	<b>3</b>	1.13 (1.39)	+
+	<b>Tender</b>	4.56 (1.44)	5.20 (1.22)	7.10***	0.50	<b>11</b>	4.07 (1.39)	5.65 (1.27)	-6.57***	1.19	<b>13</b>	4.12 (2.01)	<b>9</b>	0.35 (1.57)	○
-	Sentimental	3.79 (1.54)	4.50 (1.52)	6.99***	0.47	<b>12</b>	2.78 (1.37)	5.26 (1.52)	-7.46***	1.74	<b>8</b>	3.92 (1.94)	<b>12</b>	-1.16 (1.35)	-
+	<b>Warm-hearted</b>	4.90 (1.31)	5.45 (1.23)	6.93***	0.44	<b>13</b>	3.90 (1.30)	5.59 (1.23)	-6.86***	1.34	<b>15</b>	3.79 (1.97)	<b>15</b>	1.01 (1.50)	+
+	<b>Family-oriented</b>	4.62 (1.62)	5.29 (1.60)	6.91***	0.42	<b>14</b>	4.14 (1.39)	5.78 (1.17)	-6.57***	1.28	<b>16</b>	4.32 (2.10)	<b>6</b>	1.13 (1.62)	+
+	Affectionate	4.78 (1.32)	5.35 (1.22)	6.88***	0.46	<b>15</b>	3.83 (1.44)	5.48 (1.35)	-6.83***	1.18	<b>12</b>	3.88 (2.05)	<b>13</b>	1.01 (1.56)	+
-	Insecure	3.51 (1.47)	4.10 (1.45)	6.30***	0.41	<b>16</b>	2.80 (1.55)	4.49 (1.37)	-6.41***	1.16	<b>14</b>	3.14 (1.95)	<b>20</b>	-1.95 (1.17)	-
+	Open regarding one's own feelings	3.95 (1.76)	4.58 (1.74)	5.60***	0.36	<b>17</b>	2.59 (1.40)	5.56 (1.15)	-8.01***	2.32	<b>2</b>	4.56 (2.08)	<b>1</b>	0.03 (1.70)	○
○	Intuitive	4.69 (1.28)	4.97 (1.27)	3.42***	0.22	<b>18</b>	3.83 (1.52)	5.21 (1.35)	-5.02***	0.96	<b>20</b>	3.77 (1.99)	<b>17</b>	-0.01 (1.47)	○
+	Domestic	4.28 (1.49)	4.51 (1.49)	2.31***	0.15	<b>19</b>	3.30 (1.33)	5.54 (1.10)	-7.73***	1.84	<b>4</b>	4.23 (1.99)	<b>7</b>	-0.03 (1.59)	○
○	<b>Careful</b>	4.34 (1.41)	4.56 (1.30)	2.30*	0.17	<b>20</b>	2.88 (1.14)	5.13 (1.17)	-7.62***	1.95	<b>6</b>	3.58 (1.96)	<b>19</b>	-0.04 (1.52)	○

*Note.* Sex differences and stereotyping rank were not correlated ( $r = .18, p = .45$ ). Relevance rank was not correlated with sex differences ( $r = .24, p = .31$ ) but was marginally significantly correlated with stereotyping rank ( $r = .43, p = .06$ ); bold items remained in the final version of the questionnaire; R = Rank according to mean (difference); Val = Valence; + positive valence; - negative valence; ○ neutral valence; Val<sup>1</sup> = intended valence; Val<sup>2</sup> = rated valence.

Table S3

*Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's  $d$ ), and Ranks (R) for Sex Differences for the masculine Cognition Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1*

Masculine cognition Item	Sex differences <sup>a</sup>				
	Men	Women	Z	$d$	R
<b>To write a computer program</b>	3.11 (2.13)	1.78 (1.41)	-11.63***	-0.80	1
To operate technical devices	5.81 (1.35)	4.81 (1.51)	-11.22***	-0.68	2
<b>To understand formulas</b>	4.86 (1.64)	3.78 (1.79)	-9.41***	-0.62	3
<b>Day-to-day calculations</b>	5.32 (1.48)	4.36 (1.71)	-9.10***	-0.59	4
To recognize regularities	5.37 (1.30)	4.74 (1.37)	-7.47***	-0.47	5
<b>To solve equations</b>	4.94 (1.73)	4.11 (1.84)	-7.16***	-0.46	6
To recognize objects from a different perspective	5.37 (1.38)	4.81 (1.47)	-6.30***	-0.39	7
<b>To find a way again</b>	5.68 (1.54)	5.07 (1.74)	-6.07***	-0.36	8
<b>To follow directions</b>	5.35 (1.49)	4.80 (1.71)	-5.10***	-0.34	9
<b>To find an address for the first time</b>	5.24 (1.57)	4.73 (1.72)	-4.98***	-0.31	10

*Note.* Bold items remained in the final version of the questionnaire after factor analysis.

Table S4

*Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's  $d$ ), and Ranks (R) for Sex Differences for the feminine Cognition Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1*

Feminine cognition Item	Sex differences <sup>a</sup>				
	Men	Women	Z	$d$	R
To learn foreign languages	3.87 (1.72)	4.71 (1.62)	-7.68***	0.51	1
<b>To remember names and faces</b>	4.12 (1.74)	4.73 (1.78)	-5.51***	0.35	2
<b>To phrase a text</b>	4.94 (1.65)	5.46 (1.52)	-5.36***	0.33	3
To understand texts	5.47 (1.32)	5.85 (1.13)	-4.65***	0.32	4
To remember the exact wording of a statement	4.24 (1.57)	4.69 (1.57)	-4.52***	0.29	5
<b>To find synonyms for a word in order to avoid repetitions</b>	4.87 (1.53)	5.28 (1.42)	-4.28***	0.28	6
<b>To notice small changes</b>	4.76 (1.59)	5.14 (1.43)	-3.51***	0.26	7
<b>To explain foreign words</b>	4.66 (1.56)	4.98 (1.44)	-3.29***	0.22	8
<b>To find the right words to express certain content</b>	4.80 (1.55)	5.10 (1.52)	-3.16**	0.20	9
<b>Remembering events from your own life</b>	5.09 (1.38)	5.09 (1.42)	-0.16	0.00	10

*Note.* Bold items remained in the final version of the questionnaire after factor analysis.

ASSESSMENT OF GENDER-RELATED ATTRIBUTES – EMS 1

Table S5

*Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's d), and Ranks (R) for Sex Differences for the masculine Activities and Interests Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1*

Masculine interests Item	Sex differences <sup>a</sup>				
	Men	Women	Z	d	R
Reading car or sports magazines	3.00 (2.10)	1.46 (1.09)	-14.59***	-1.04	1
Working on cars/motorbikes, etc.	2.83 (1.95)	1.45 (1.08)	-14.43***	-0.97	2
Interest in fast vehicles (cars, motorbikes, quads)	3.89 (2.12)	2.29 (1.79)	-12.40***	-0.84	3
Computer and video games	4.99 (1.93)	3.36 (2.06)	-12.03***	-0.81	4
<b>Watching sports on TV (boxing, Formula 1, ball games...)</b>	3.93 (2.31)	2.22 (1.68)	-11.97***	-0.90	5
Mountain biking	4.04 (1.96)	2.70 (1.88)	-10.64***	-0.70	6
<b>Watching action movies</b>	4.95 (1.63)	3.67 (1.91)	-10.43***	-0.70	7
<b>Doing certain sports (soccer, basketball, hand ball, etc.)</b>	4.84 (2.09)	3.43 (2.09)	-10.16***	-0.68	8
<b>Driving go-carts</b>	4.05 (2.15)	2.75 (1.88)	-9.64***	-0.66	9
<b>Drinking beer</b>	4.79 (2.04)	3.49 (2.17)	-9.26***	-0.61	10
<b>Paintball</b>	4.05 (2.23)	2.79 (1.93)	-8.90***	-0.62	11
Listening to "hard" music	4.28 (2.01)	3.44 (2.17)	-6.21***	-0.40	12
<b>Playing cards (poker)</b>	4.37 (1.88)	3.60 (1.96)	-6.15***	-0.40	13
<b>Gym (weightlifting)</b>	3.68 (2.16)	2.96 (1.98)	-5.34***	-0.35	14
Driving a car	5.02 (1.84)	4.54 (1.99)	-3.83***	-0.25	15

*Note.* Bold items remained in the final version of the questionnaire after factor analysis.

Table S6

*Results of Ratings in Means (SD), Z-Scores, Effect Sizes (Cohen's  $d$ ), and Ranks (R) for Sex Differences for the feminine Activities and Interests Items from the shorter Version of the GERAS after first Reduction of the Item-pool from Study 1*

Feminine interests Item	Sex differences <sup>a</sup>				
	Men	Women	Z	$d$	R
Beauty products and make-up	1.44 (1.06)	4.10 (1.93)	-20.95***	1.56	1
<b>Shopping</b>	3.08 (1.71)	4.86 (1.80)	-14.33***	1.01	2
Decorating an apartment	2.92 (1.66)	4.63 (1.79)	-13.95***	0.98	3
Baking	3.35 (1.87)	5.05 (1.74)	-13.53***	0.95	4
Crocheting/ sewing/ knitting	1.69 (1.24)	3.39 (2.15)	-13.30***	0.89	5
Reading novels	3.77 (1.97)	5.32 (1.78)	-12.35***	0.84	6
<b>Watching a romantic movie</b>	2.76 (1.56)	4.18 (1.88)	-11.73***	0.80	7
<b>Talking on the phone with a friend</b>	3.08 (1.76)	4.47 (1.92)	-10.95***	0.74	8
<b>Yoga</b>	2.03 (1.56)	3.35 (2.02)	-10.82***	0.70	9
<b>Rhythmic gymnastics</b>	1.61 (1.18)	2.76 (1.93)	-10.24***	0.67	10
<b>To gossip</b>	3.61 (1.76)	4.78 (1.71)	-10.21***	0.68	11
<b>Dancing</b> (classic standard dances, ballet, Latin, free dance, etc.)	2.72 (1.89)	4.10 (2.19)	-9.79***	0.66	12
<b>Going for a walk</b>	4.26 (1.72)	5.25 (1.55)	-9.25***	0.62	13
Spa vacation	3.14 (1.92)	4.21 (2.09)	-8.07***	0.53	14
Drawing/painting/pottery	2.90 (1.82)	3.86 (2.08)	-7.26***	0.48	15

*Note.* Bold items remained in the final version of the questionnaire after factor analysis.

<sup>a</sup>  $N = 1,142$ ; 364 men ( $M_{\text{age}} = 25.29$ ,  $SD = 5.75$ ) and 778 women ( $M_{\text{age}} = 24.37$ ;  $SD = 4.97$ ) matched on age, education, and socioeconomic status were selected to assess sex differences in each item; only items with the largest sex differences were included in the PP version of the GERAS.

<sup>b</sup>  $N = 98$  participants (30 men,  $M_{\text{age}} = 25.83$ ;  $SD = 7.48$ ; 68 women,  $M_{\text{age}} = 25.96$ ;  $SD = 8.93$ ) rated how strongly they associated these items with a stereotypical man on the one hand and a stereotypical woman on the other hand on a 7-point Likert scale; only items with the largest differences between the ratings for a stereotypical man and a stereotypical woman were included in the PP version of the GERAS.

<sup>c</sup>  $N = 105$  participants (30 men,  $M_{\text{age}} = 24.47$ ;  $SD = 5.30$ ; 75 women  $M_{\text{age}} = 23.79$ ;  $SD = 4.39$  years) rated personality items for their relevance for determining gender role orientation on a 7-point Likert scale; only the items with the highest relevance ratings were included in the PP version of the GERAS.

<sup>d</sup> The same subgroup of participants as described in <sup>c</sup> rated the social desirability of each item on a 7-point Likert scale; differences in social desirability ratings and proposed valence are marked in yellow; positive, negative, and neutral traits were included in the scale to account for response tendencies in terms of social desirability.