Sex Roles (2006) 55:247–258
DOI 10.1007/s11199-006-9077-0

ORIGINAL ARTICLE

Healthy Women, Healthy Men, and Healthy Adults: An Evaluation of Gender Role Stereotypes in the Twenty-first Century

Susan Rachael Seem · M. Diane Clark

Published online: 8 November 2006

Abstract An important question often asked when counselors-in-training read textbook discussion of gender role stereotypes, especially of older work such as the classic study by the Brovermans and their colleagues, is "Haven't these biases been eliminated or at least reduced?" The current study was designed to replicate the work of the Brovermans and their colleagues to answer that specific question and to determine how current counselors-in-training perceive healthy adult women, healthy adult men, and healthy adults. As in the prior research, initial ratings of the social desirability of traditional gender role stereotypes were conducted, and the findings showed many similarities to past research. That investigation was followed by a modified Stereotype Questionnaire, based on the original work of Rosenkrantz, Vogel, Bee, I. Broverman, and D. M. Broverman (1968). Healthy adult women were found to be significantly different from healthy adult men as well as from healthy adults. In addition, the results suggest that there have been changes in counselors' perceptions of healthy adults. Counselors-in-training were found to hold two standards for mental health—one for women and another for men.

Keywords Gender roles · Stereotypes · Counselors-in-training

S. R. Seem (✉)
Department of Counselor Education,
State University of New York College at Brockport,
350 New Campus Drive, Brockport,
NY 14420, USA
e-mail: sseem@brockport.edu

M. D. Clark
Gallaudet University, Washington, DC, USA

Relationships among self-concept, psychological health, and gender stereotypes emerged in the work of Rosenkrantz, Vogel, Bee, I. Broverman, & D. M. Broverman, (1968). In order to examine these relationships, Rosenkrantz et al. investigated first the social desirability of traits used to describe men and women and then developed the Stereotype Questionnaire. The initial listing of socially desirable items was completed by having two classes of undergraduate students write down the characteristics that they believed differentiated men and women. Any item listed more than once was included in the Stereotype Questionnaire, which consisted of 122 items that were arranged along a Likert scale in a bipolar format (e.g., "not at all aggressive"–"very aggressive"). An interesting component of this 7-point Likert scale is that there were ten potential points between each of the Likert anchor-numbers for a total of 60 possible points. Rosenkrantz et al. then tested the Stereotype Questionnaire with a second independent sample of undergraduates to determine which pole of each of the 122 bipolar items represented the socially desirable one.

The second study of Rosenkrantz et al. (1968) focused on determining the gender role stereotypes for men and for women and comparing those stereotypes to the ideal of a healthy adult. Their procedure asked all participants to "imagine that you are going to meet a person for the first time and the only thing you know in advance is that this person is an adult (man)" (p. 288)/adult woman. A within group design was used and participants were asked to rate each of the items to the extent that it characterized: (a) a healthy adult man, (b) a healthy adult woman, and (c) themselves (given the assumption that the later rating would represent a healthy adult, sex unspecified). Presentation of the healthy adult man and healthy adult woman instructions was counterbalanced and was always followed by the rating of a healthy adult. The ratings were used to determine which traits represented gender

role stereotypes by comparing the characteristics for a healthy adult man and for a healthy adult women to the characteristics for a healthy adult. Results indicated that college students, regardless of sex, agreed on the gender role stereotypes of men and women. They additionally concluded that the self-concepts of male and female participants followed these traditional gender role stereotypes meaning that women "presumably,... also hold negative values of their worth relative to men" (p. 293). Finally, they found a larger number of socially desirable characteristics and behaviors were stereotypically connected with masculinity than with femininity.

In an extension of the study of Rosenkrantz et al. (1968), Broverman and colleagues, (I. K. Broverman, D. M. Broverman, Clarkson, Rosenkrantz, & Vogel, 1970; I. K. Broverman, Vogel, D. M. Broverman, Clarkson, & Rosenkrantz, 1972) investigated female and male clinicians' (i.e., psychologists, psychiatrists, and social workers) judgments of the mental health of women and men. As in the earlier work of Rosenkrantz et al., the Stereotype Questionnaire was used, but this time with a between groups design. In their description of the instrument, Broverman et al. (1970) reported use of the questionnaire of Rosenkrantz et al. with 122 bipolar items. Their example in the 1970 article, however, did not include a Likert scale; instead it was posed as a forced choice format ("not at all aggressive"–"aggressive").

Broverman et al. (1970) hypothesized that (a) clinical judgments of the characteristics of a healthy, mature individual would differ as a function of the sex of the person judged and that (b) behavioral characteristics that were regarded as healthy for an adult, sex unspecified, would be more often regarded as healthy for men than for women (i.e., following cultural stereotypes of gender differences). Two types of scores were calculated: health scores and agreement scores. The health scores were based upon the assumption that traits selected for healthy adults would reflect the definition of mental health for all individuals. These health scores were determined by the pole selected by 75% or more of the participants who were given the healthy adult instructions for each of the 122 items. The masculinity agreement score was based on the percentage of participants who selected the socially desirable pole with the healthy adult man instructions and the femininity agreement score was based on the percentage who selected the socially desirable pole with the healthy adult woman instructions.

The results of Broverman et al. (1970) paralleled the findings of Rosenkrantz et al. (1968) and showed that clinicians' judgments of adult men's mental health did not differ significantly from their judgments of healthy adults, sex unspecified, whereas their judgments of adult women's mental health did differ significantly from judgments of healthy adult men and healthy adults sex unspecified. This determination was based upon the large overlap among health scores and masculinity agreement scores in contrast to the limited overlap among health scores and femininity agreement scores. These differences corresponded with cultural stereotypes of men and women that were widely held during that time. Healthy adult women were described as different from both healthy adult men and healthy adults in that they were more "submissive," "less independent," "less aggressive," "less competitive," "more easily influenced," "more emotional," and "less objective." The authors concluded that a double standard of mental health existed for women, in that for a woman to be seen as mentally healthy she must be feminine and not adult-like (i.e., not like a man). Both female and male clinicians implicitly supported this standard. Furthermore, the authors discovered that clinicians when given instructions to "describe healthy, socially competent women" (p. 2) used fewer traits than when they were given instructions to" describe healthy, socially competent men" (p. 2). The traits found for healthy women were often viewed as less socially desirable than the traits listed for healthy men.

In an attempt to evaluate any changes in mental health professionals' gender role stereotypes and clinical judgments, Phillips and Gilroy (1985) used a shortened version of the Likert scale Stereotype Questionnaire created by Rosenkrantz et al. (1968) in a between groups design. As in the Broverman et al. (1970) study, the participants were psychiatrists, social workers, and psychologists. Similar to Broverman et al., they found no significant differences in ratings of health for adult men, adult women, and adults, sex unspecified, that were related to the clinician's sex. In contrast to the results of Broverman et al., Philips and Gilroy found high levels of overlap when they compared the health scores to both the masculinity agreement scores and the femininity agreement scores. Philips and Gilroy challenged the conclusion of Broverman et al. regarding a double standard of mental health, and argued that the conclusion of Broverman et al. was based on a statistical artifact related to their use of a forced choice format. Furthermore, Philips and Gilroy found no significant relationship of the social desirability of traits and traditional gender role stereotypes, and they postulated that most traits were viewed as socially desirable for both sexes. Phillips and Gilroy concluded that either the Broverman et al. results were due to this statistical artifact or that their own results were due to the fact that progress had been made by clinicians in the reduction of their gender role stereotyping in clinical judgments.

In another replication of Broverman et al. (1970), Widiger and Settle (1987) addressed methodological concerns in the classic design by focusing on the imbalanced ratio of traditionally masculine valued to traditionally feminine valued items as well as the use of bipolar adjectives. Widiger and Settle argued that feminine traits are not the opposite of masculine traits; they found that if

healthy adult men were rated high on an item, it did not necessarily mean that healthy adult women were not rated relatively high on that same item. Widiger and Settle, therefore, focused on developing unipolar items—i.e., each item was paired with its negative, not its opposite (e.g., "very independent" versus "not at all independent" rather than "independent" versus "dependent"). Furthermore, 72 additional items were created and rated for social desirability by college students in order to correct for the imbalance of traditionally masculine valued items in Broverman et al., who had a ratio of 71% traditionally masculine valued versus 39% traditionally feminine valued items.

Given these changes, Widiger and Settle (1987) contended that the conclusions of the Broverman et al. (1970) study were a result of a statistical artifact that was due to the ratio of traditionally masculine valued to traditionally feminine valued items on the dependent measure. They argued that because most of the socially desirable poles in the Broverman et al. study "concerned a stereotypically masculine characteristic, it is not surprising that the mean masculinity health score was close to the adult health score and greater than the femininity health score" (p. 464). The results, they reasoned, depended upon how the analysis was done. By altering the ratio of traditionally masculine valued to traditionally feminine valued items in three analyses (i.e., first analysis: 27 traditionally feminine valued items versus 11 traditionally masculine valued items, second analysis: 27 traditionally masculine valued items versus 11 female-valued items, and third analysis: an equal number of female- and traditionally masculine valued items), the authors found gender bias against women, gender bias against men, and no gender bias, respectively, based solely on statistical measures.

In a recent partial replication of the classic study by Rosenkrantz et al. (1968), Nesbitt and Penn (2000) surveyed community college students using Stereotype Questionnaire of Rosenkrantz et al. in order to investigate whether gender role stereotypes had changed in the 30 plus years since the original study. Although gendered traits such as expressing and experiencing affect still differentiated healthy adult women and healthy adult men, gender role stereotypes had changed in that traits such as "logical," "ambitious," "direct," and "the ability to separate ideas from feelings" no longer differentiated women from men.

In conclusion, some of the prior research on the impact of gender role stereotyping on clinical judgement showed bias against women, whereas the results of other studies suggested that the dependent measure had actually elicited the bias. Thus, the results are inconclusive. Despite criticism of the methodology of the Broverman et al. (1970) investigation (Phillips & Gilroy, 1985; Widiger & Settle, 1987) and the fact that it was conducted over 30 years ago, that study is still cited in college textbooks (see, for example, Crawford & Unger, 2000; Lerner, 2002). Students consistently ask whether the conclusions of Broverman et al. are still true. Given this fact, it seems important to reexamine gender role stereotyping by future counselors in the twenty-first century.

Objectives

In the present research study, as in the prior research discussed above, two studies were conducted: one to assess social desirability, and the other to investigate gender role stereotypes with the Stereotype Questionnaire. The objective of Study 1 was to evaluate the current social desirability of the items to be used in Study 2 on the Stereotype Questionnaire. Therefore in Study 1, participants were asked to pick the socially desirable pole of each item. Study 2 had two objectives. The first objective was to evaluate current gender role stereotypes held by counselors-in-training regarding healthy adult women, healthy adult men, and healthy adults, sex unspecified, using the Stereotype Questionnaire. The second objective was to determine whether gender role stereotypes for men and women have changed since the studies of Broverman et al. (1970, 1972).

Study 1

Method

Participants The participants in the social desirability study were 89 undergraduate, introductory psychology students enrolled in a comprehensive university in the northeastern United States. They were 47 women and 42 men with a mean age of 21.8 years (range—18 to 29; mode = 18). Ninety-eight percent of the participants were European Americans. The other 2% included one African American and one Asian American. These percentages represent the demographics of the student population at this university. Participants were given extra credit for their participation.

Measure A Social Desirability Questionnaire was created that consisted of the 62 unipolar items later used in the Stereotype Questionnaire. This questionnaire was more closely aligned to the Widiger and Settle (1987) questionnaire than to the Broverman et al. (1970, 1972) questionnaire, which was in a bipolar format. Several new items were developed for the present study to correct for the 14 bipolar items that remained in Widiger and Settle's version. Therefore, for the current study, each of these 14 items was split into two unipolar items for a total of 28 new items. For example, "strong" versus "weak" became "very strong" versus "not at all strong" and "very weak" versus "not at all weak." Pilot testing eliminated seven of the new items

because participants did not indicate a socially desirable characteristic for those items. The Stereotype Questionnaire for this study consisted of 62 unipolar pairs (see Table 1 for a complete list of all unipolar pairs). The items were presented as pairs in two columns, and participants were directed to circle one item from each pair that represented the trait that was more socially desirable. Items were counterbalanced so that one-half of the historically, socially desirable items were presented in the left-hand column, whereas the other one-half were presented in the right-hand column.

Analytic plan The social desirability of each of the 62 unipolar items was determined by counting the number of times each description (i.e., the left hand side—"very sympathetic" versus the right hand side—"not at all sympathetic") was selected. This social desirability determination was summed across participants' responses for each item and recorded as a percentage. A z score was used as the criterion to determine if participants noted a socially desirable pole between the two descriptions for each item (i.e., if one description was selected significantly more often than the other one). In order to declare one description of an item as socially desirable, a z score above 1.65 was required. This z value was established by the z formula (N observed$-N$ expected/the square root of N expected) to obtain the required percent. A percentage of 60.03% was determined as follows (N observed $= 89/2 + 1.65$[square root of 89] $=$ 60.03) and rounded up to 61% such that $p<0.05$ (two-tailed).

Results and Discussion

Five of the 62 pairs did not have social desirability scores above 61% for either description of the item. These items included: "not at all uncomfortable when people express emotions," "not at all dependent," "acts on logic rather than feelings," "very strong need for security," and "not at all passive." (See Table 1). As can be seen in Table 1, the remaining 57 pairs had social desirability scores above 61%.

Participants viewed in traditional ways all but one of the items from past research. "Very subjective" versus "not at all subjective" differed from past research in that historically, the socially desirable masculine pole of "not at all subjective" was replaced by the new socially desirable feminine pole "of very subjective." All of the remaining traits continued to be viewed as they have been historically.

The items that did not obtain 61% social desirability percentages include two that were historically feminine items ("not at all uncomfortable when people express emotions" and "very strong need for security") and two that were historically masculine items ("not at all dependent" and "acts on logic rather than feelings"). The remaining item that did not obtain a 61% score was a new item ("very passive") that in prior research had been paired with "active." It appears that "passive" is not seen as a socially desirable trait.

In general, the results show that all participants had a clear understanding of cultural values regarding what is considered socially desirable by our society (most items received socially desirable scores above 80%). These items are highly similar to those found by Rosenkrantz et al. (1968) and Broverman et al. (1970, 1972).

Study 2

Method

Participants The participants in Study 2 were 121 students enrolled in two masters' level counseling programs in the northeastern United States. The mean age was 28.4 years (range—21 to 48; mode=23). There were 65 (54%) women and 56 (46%) men; 112 were European Americans, six were African Americans, and three were Asian Americans. This represents the demographics of the student bodies in both universities. Students were enrolled in four degree programs: masters of science in education (61%), masters of science (25%), masters of arts (10%), and certificate of advanced study (4%). These students were recruited from classes taught by the authors as well as by other members of their departments. Independent graduate students (who were not in the authors' classes) gave the surveys to the students who chose to participate; all surveys were returned to the respective university department's office.

Measure Items for the Stereotype Questionnaire were developed in Study 1. See Table 1 for a complete list of the items. The items were arranged on a Likert scale in a unipolar format (e.g., "not at all aggressive"–"very aggressive"), and anchored on a 1 to 7 scale as in Rosenkrantz et al. (1968).

Procedure Each participant received a package that included the three Stereotype Questionnaires, as in the Rosenkrantz et al. (1968) within group design. Directions specified that participants were to complete each survey in the order presented. In addition, participants were informed that, although each survey was the same, it would be preceded by a different set of instructions. They were warned to read each new set of instructions carefully before completing the questionnaire. All participants were given the following instructions: "Imagine that you are going to meet a person for the first time and the only thing you know in advance is that the person is a woman (man/adult). On the following questionnaire, think of a normal man (woman/adult), and then indicate on each item the pole to which a mature, healthy,

Table 1 Results of the social desirability questionnaire with socially desirable items presented in the left column.

Social desirability scores			
Item	Percent agreement	Item	Percent agreement
Unipolar items			
Traditionally masculine items			
Controls emotion*	90	Does not control emotions	10
Decisive*	84	Indecisive	13
Very worldly*	85	Not at all worldly	13
Very firm*	87	Not at all firm	13
Controls self under stress*	92	Becomes upset under stress	8
Very direct*	90	Not at all direct	10
Very active*	96	Not at all active	4
Very adventurous*	94	Not at all adventurous	6
Not at all sneaky*♦	71	Very sneaky	29
Heroic*	88	Not heroic	11
Knows the way of the world*	96	Does not know the way of the world	4
Very Strong*	95	Not at all strong	4
Easily able to separate feelings from ideas*	83	Unable to separate feelings from ideas	16
Very independent*	98	Not at all independent	2
Relies on self*	85	Relies on others	11
Enjoys a challenge very much*	93	Does not enjoy a challenge very much	7
Not at all weak*♦	94	Very weak	6
Very industrious*	93	Not at all industrious	7
Can make decisions easily*	92	Has difficulty making decisions	8
Very daring*	95	Not at all daring	4
Not at all afraid to take risks*	93	Very afraid to take risks	7
Very brave*	92	Not at all brave	8
Not at all excitable in a minor crisis*	72	Very excitable in a minor crisis	28
Has a strong will*	98	Does not have a strong will	2
Very authoritative*	67	Not at all authoritative	32
Very competitive*	84	Not at all competitive	16
Very objective*♦	73	Not at all objective	26
Unipolar pairs			
Traditionally feminine items			
Very concerned about others*	94	Not at all concerned about others	6
Very neat in habits*	92	Not at all neat in habits	8
Enjoys art and literature very much*	81	Does not enjoy art and literature very much	18
Very sensitive*	87	Not at all sensitive	13
Very talkative*	87	Not at all talkative	13
Very gentle*	92	Not at all gentle	8
Very careful*	96	Very careless	4
Very affectionate*	87	Not at all affectionate	13
Not at all harsh*♦	83	Very harsh	17
Very idealistic*	78	Not at all idealistic	21
Very sympathetic*	90	Not at all sympathetic	10
Very sentimental*	90	Not at all sentimental	10
Very charitable*	92	Not at all charitable	8
Very understanding of others*	98	Not at all understanding of others	2
Very creative*	94	Not at all creative	6
Very emotional*	65	Not at all emotional	34
Very warm in relations with others*	95	Not at all warm in relations with others	5
Very aware of feelings of others*	95	Not at all aware of feelings of others	5
Very considerate of others*	96	Very inconsiderate of others	4
Very interested in own appearance*	83	Not at all interested in own appearance	16
Very subjective*	72	Not at all subjective	27
Very compassionate*	95	Not at all compassionate	5
Not at all cold*♦	96	Very cold	2

Table 1 Continued

Social desirability scores			
Item	Percent agreement	Item	Percent agreement
Very soothing*	94	Not at all soothing	5
Very home oriented*♦	77	Not at all home oriented	23
Not at all rough*♦	71	Very rough	28
Easily expresses tender feelings*	88	Does not express tender feelings at all	12
Very soft*♦	72	Not at all soft	28
Very loving*	98	Not at all loving	2
Very forgiving*	95	Not at all forgiving	5
Unipolar pairs			
Non-significant agreement scores			
Not at all uncomfortable when people express emotion (F)	55	Very uncomfortable when people express emotions	45
Not at all dependent (M)	52	Very dependent	48
Acts on logic rather than feeling (M)	60	Acts on feeling rather than logic	39
Very strong need for security (F)	49	Very little need for security	51
Very passive (F)	49	Not at all passive	51

*Socially desirable pole; ♦ new item based on the non-socially desirable pole in prior research
F Historically/traditionally feminine item, *M* historically/traditionally masculine item

socially competent woman (man/adult) would be closer. Indicate your degree of conviction by placing a slash through the point on the scale at which you feel this man (woman/adult) would best be described by the trait under consideration." The order of presentation of the healthy adult woman and a healthy adult man Stereotype Questionnaires was counterbalanced, and the healthy adult always followed these.

Analytic plan The analytical plan included three types of analyses. First, we calculated item agreement scores. Next, binary, stereotypic items were determined. Then, continuous, total, gender stereotypic scores were calculated and used in a repeated measure MANOVA with paired *t*-test post-hocs. Each of these coding schemes are described below.

The 7-point Likert scale used in the Stereotype Questionnaire to present the unipolar items was analyzed as a binary code to create item agreement scores. Because no neutral (i.e., a score of 4) scores were obtained, scores below the neutral rating of 4 (1 through 3) were recoded as 1. Scores above the neutral rating of 4 (5 through 7) were recoded as 2. Individual item agreement scores then were calculated by determining the percentage of participants who selected the socially desirable item within each of the 62 pairs. These poles were labeled either traditionally feminine or traditionally masculine, based on prior research. Therefore, six sets of item agreement scores were calculated for each socially desirable item, one for female participants and another for male participants, on each of the three Stereotype Questionnaires—healthy adult woman, healthy adult man, and healthy adult. Specifically, each of these six sets had 62 item agreement scores for a total of 372 item agreement scores across the six sets.

Binary stereotypic items were determined by combining men's and women's item agreement scores on each of the 62 pairs across the three Stereotype Questionnaires. Then, groups of stereotypic items were developed by organizing the item agreement scores discussed above into lists, one for men, one for women, and one for adults. Items were selected for inclusion if 70% or more of the participants had selected the socially desirable pole for that item. Seventy percent was used rather than the 75% in earlier work as examination of the current data showed this percentage to be a natural break point and this cut-score was more stringent than the 61% used in Study 1. (See Study 1 for a discussion related to the determination of percentage cut-offs using *z* scores.)

Then, the continuous data originally collected for each item was selected for all stereotypic items. The participants' responses on these items were then averaged together to create six total gender stereotypic scores, two each for a healthy adult woman (i.e., one each for men and women), two each for a healthy adult man, and two each for a healthy adult. Scores like these total gender stereotypic scores were calculated by both Broverman et al. (1970, 1972) and Widiger and Settle (1987).

The six total gender stereotypic scores of the healthy adult woman, healthy adult man, and healthy adult, described above, were analyzed in a repeated measure MANOVA. Sex was used as a between participant factor. Planned comparisons

with paired *t*-tests were performed to compare a healthy adult woman to a healthy adult, a healthy adult man to a healthy adult, and a healthy adult woman to a healthy adult man.

Results

Item agreement scores showed that there were 35 traditionally feminine items and 27 traditionally masculine items across the Stereotype Questionnaires. Item agreement scores for a healthy adult woman were generally above 50% on both the traditionally feminine and masculine items with a few notable exceptions. Within this view of a healthy adult woman, female participants' item agreement scores were less than 50% on only two traditionally feminine items—"subjective" and "very strong need for security." Here, female participants did not view these items as descriptive of a healthy adult woman. Male participants tended to report higher item agreement scores on these traditionally feminine items than did their female counterparts. One additional traditionally feminine item agreement score for men was below 50% ("not at all weak"), whereas the comparable female participants' item agreement score was higher (60%). In terms of traditionally masculine items, female participants had 12 item agreement scores less than 50% for a healthy adult woman; whereas male participants had 14 items (see Table 2). Therefore, participants, regardless of their sex, endorsed many traditionally masculine items at relatively high levels (above 50%) as characteristic of a healthy adult woman (15 by female and 13 by male participants; see Table 2—healthy adult women—masculine pole socially desirable).

Examination of the data for a healthy adult man revealed that participants reported higher overall scores on the traditionally masculine items than on the traditionally feminine items (see Table 2). On the traditionally masculine items, only four items had agreement scores below 50%, and only female participants reported three of those scores. In contrast to these relatively high traditionally masculine item agreement scores, only five traditionally feminine items had agreement scores above 50% for a healthy adult man. These items included "idealistic," "not at all weak," "interested in own appearance," "loving," and "acts on feelings rather than logic." Participants' sex differences were also reflected in higher item agreement scores by female participants than by male participants on traditionally feminine items (i.e., only women reported item agreement scores above 70% when they evaluated a healthy adult man on traditionally feminine items).

The healthy adult item agreement scores tended to be more equivalently endorsed by female and male participants for both traditionally feminine and traditionally masculine items. Items were more likely to have agreement scores below 40% if they were traditionally feminine items. Seven items were reported below 40% by female participants and seven items by male participants; see Table 2. On the other hand, items on the healthy adult questionnaire were more likely to have item agreement scores above 60% if they were traditionally masculine items. Here 11 items were reported by female participants and 15 by male participants. Male participants had high item agreement scores on these traditionally masculine items. Their female counterparts also had high item agreement scores on these items, except for the items "competitive" (52%), "adventurous" (52%), and "firm" (52%).

Stereotypic items The perception of a healthy adult woman included 21 stereotypic items (see Table 3 for a complete list). All but three of these items were from the traditionally feminine item pool. The three items not from the traditionally feminine item pool included "strong," "independent," and "enjoys a challenge." Therefore, the healthy adult woman description included 18 traditionally feminine items as well as three traditionally masculine items.

The healthy adult man description included 15 stereotypic items (see Table 3 for a complete list). All of these items were from the traditionally masculine pool. Here the healthy adult man had fewer stereotypic items than did the healthy adult woman.

The healthy adult description included five stereotypic items that reached the 70% agreement criteria (see Table 3 for a complete list). One of these items was from the traditionally feminine pool ("understanding of others"), and the remaining four items were from the traditionally masculine pool. The healthy adult had many fewer endorsed items than did either the healthy adult woman or the healthy adult man.

MANOVA and paired t-tests A repeated measures MAN OVA was performed to compare the total gender stereotypic scores. Participant sex was the between subject factor. Both female participants (SD=0.26) and male participants (SD= 0.50) had means of 3.97 on the total gender stereotypic scores for the healthy adult women. For the healthy adult man the total gender stereotypic scores showed sex differences; female participants had a mean score of 4.21 (SD= 0.25) and male participants had a mean score of 4.34 (SD= 0.49). Total gender stereotypic scores for the healthy adult also were the same (mean score of 4.25) for female participants (SD=0.34) and male participants (SD=0.37).

There was a significant effect for total gender stereotypic scores, F (1, 86)=30.81, p 0.0001, but the interaction between these scores and participant sex was not significant, F (1, 86)=0.03, p 0.87. To evaluate the main effect of the total gender stereotypic scores, paired *t*-tests were performed. These analyses showed significant differences

Table 2 Individual percent item agreement scores for the three Stereotype Questionnaires presented separately by participant sex.

	Healthy adult woman		Healthy adult man		Healthy adult	
Participant sex	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)
Feminine pole—traditionally socially desirable						
Idealistic	64	62	52	59	59	55
Sympathetic	74	79	31	26	59	55
Sentimental	71	75	24	32	44	50
Charitable	74	74	46	44	63	53
Understanding of others	83	81	41	42	71	71
Creative	69	73	43	44	57	62
Not at all weak	60	47	75	66	59	55
Emotional	64	69	13	17	29	30
Warm in relations with others	84	87	40	42	60	60
Considerate of others	81	87	48	51	72	57
Interested in own appearance	76	94	57	59	64	78
Subjective	44	53	35	42	35	31
Concerned about others	86	77	41	51	70	64
Neat in habits	67	84	40	30	47	54
Enjoys art and literature	65	80	34	33	45	51
Sensitive	81	90	26	35	54	63
Talkative	68	70	32	35	48	64
Gentle	75	82	39	32	48	47
Careful	66	72	40	46	50	69
Affectionate	80	93	45	40	52	52
Not at all sneaky	48	49	34	36	41	46
Compassionate	87	87	43	43	69	61
Soothing	81	82	44	40	59	50
Home oriented	52	54	34	25	53	48
Easily expresses tender feelings	56	66	32	27	44	36
Soft	50	64	23	14	32	29
Loving	89	85	55	61	59	61
Forgiving	69	74	43	47	47	58
Acts on feelings rather than logic	49	66	72	62	17	13
Not at all uncomfortable when people express emotions	67	58	55	54	44	35
Aware of the feelings of others	86	80	31	30	60	57
Not at all cold	77	72	45	38	61	44
Not at all harsh	60	54	23	15	30	25
Not at all rough	52	54	21	7	37	26
Very strong need for security	44	54	28	38	39	43
Masculine pole—traditionally socially desirable						
Knows the way of the world	59	76	73	72	70	76
Strong	79	67	72	85	75	76
Easily able to separate feelings from ideas	60	49	39	42	55	52
Independent	72	76	84	85	75	73
Relies on self	49	53	69	74	57	61
Enjoys a challenge	76	75	81	83	73	75
Industrious	60	53	79	76	65	63
Can make decisions easily	48	44	75	63	66	66
Daring	32	41	75	79	54	63
Controls emotions	29	26	65	78	38	44
Decisive	52	47	74	74	70	61
Worldly	52	44	62	63	66	65
Firm	43	53	73	76	52	62
Controls self under stress	44	36	66	64	58	58
Direct	54	58	77	78	65	64
Active	59	68	68	77	54	36
Adventurous	51	56	72	87	52	64
Heroic	54	40	62	62	62	51

Table 2 Continued

Participant sex	Healthy adult woman		Healthy adult man		Healthy adult	
	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)
Not at all afraid to take risks	44	40	78	68	50	58
Brave	57	46	73	74	49	58
Not at all excitable in a minor crisis	33	33	60	61	41	48
Has a strong will	67	71	79	78	67	68
Authoritative	48	33	71	70	45	42
Competitive	53	48	82	85	52	62
Objective	47	43	48	60	42	51
Not at all dependent	36	47	23	60	48	46
Not at all passive	35	31	65	52	8	13

between a healthy adult woman and a healthy adult, $t=-5.72$, $df=92$, p 0.0001, and between a healthy adult woman and a healthy adult man, $t=-5.22$, $df=89$, p 0.0001. There was no significant difference between a healthy adult man and a healthy adult, $t=0.51$, $df=93$, p 0.61.

Discussion

Several interesting results emerged when perceptions of a healthy adult woman, a healthy adult man, and a healthy adult were compared. First, individuals endorsed more traits for a healthy adult woman than was true in the past. The increase in the number of traits used to describe a healthy adult woman suggested that current gender role stereotypes have changed, but maybe not for the better. Now women are expected not only to be nice and nurturing (traditional expectations) but also to demonstrate traits in the traditionally masculine area of competency (De Lisi & Soundranayagam, 1990; Nesbitt & Penn, 2000).

In addition, in this study, a healthy adult woman had more endorsed traits than either a healthy adult man or a healthy adult. Men's item agreement scores for traditionally feminine items on the healthy adult women's stereotypic items were generally higher than those reported by women. Men continue to appear to expect women to be more traditionally feminine than women themselves believe is appropriate. It is interesting that both sexes agree that descriptions of a healthy adult woman include behaviors that are considered traditionally masculine. For example, high consistency occurred between the sexes on the traits "independent" and "enjoys a challenge." However, women's item agreement score on "strong" was much higher than that reported by men.

In contrast to the changes in gender role stereotypes for a healthy adult woman, perceptions of a healthy adult man have showed little change since the early work of Rosenkrantz et al. (1968). Men are still expected to display traditionally masculine characteristics and behaviors. In contrast to the healthy adult woman, who now is described in both traditionally feminine and masculine terms, the healthy adult man is still perceived solely in traditionally masculine terms. The soft, sensitive man may play well in the movies—but our findings suggest that counselors-in-training would find such a man less than healthy.

Most of the few traits used to describe a healthy adult continue to be traditionally masculine. Given that most item agreement scores were around the 50% mark, it is possible that a healthy adult is perceived to be one who is neither strongly feminine nor strongly masculine, but rather more androgynous. On the other hand, it may be that there is no clear consensus regarding how to describe a healthy adult. Future researchers need to clarify these results to determine whether or not it is possible to conceive of a healthy adult outside of the constraints of traditional gender roles or whether this finding was related to participant bias. It is possible that given the within subjects design of the study, that participants had "guessed" what responses were desired and thereby formed their answers based on these assumptions. Future research could check these issues by having an independent sample complete only the section for a healthy adult, sex unspecified.

The most important finding is that a healthy adult woman continues to be significantly different than a healthy adult man and a healthy adult, sex unspecified. Yet again, a healthy adult man was not found to be significantly different from a healthy adult, sex unspecified. Stereotypic items between a healthy adult woman and a healthy adult man overlap only on the three traditionally masculine items of "strong," "independent," and "enjoys a challenge." Otherwise, the remaining 17 stereotypic items for a healthy adult woman continued to be traditionally feminine, whereas the remaining 12 items for a healthy adult man continued to be traditionally masculine. A healthy adult woman also continued to be significantly different from a healthy adult whereas a healthy adult man was not found to be

Table 3 Stereotypic items included in the healthy adult woman, healthy adult man and healthy adult stereotypic scores.

Healthy adult woman—stereotypic items		Healthy adult man—stereotypic items	
Item	Percent	Item	Percent
Sympathetic	77	Knows the way of the world	72
Sentimental	70	Strong	77
Charitable	78	Independent	86
Understanding of others	83	Enjoys a challenge	83
Warm in relations to others	86	Industrious	75
Considerate of others	86	Daring	76
Interested in own appearance	83	Decisive	71
Concerned about others	81	Firm	75
Neat in habits	75	Direct	79
Enjoys art and literature	72	Adventurous	81
Sensitive	83	Not at all afraid to take risks	74
Gentle	78	Brave	74
Affectionate	85	Has a strong will	79
Compassionate	87	Authoritative	70
Soothing	82	Competitive	83
Loving	87		
Independent	71		
Enjoys a challenge	73		
Aware of the feelings of others	78		
Strong	71		
Not at all cold	75		
Healthy adult—stereotypic items			
Item	Percent		
Understanding of others	71		
Knows the way of the world	73		
Strong	76		
Independent	76		
Enjoys a challenge	74		

significantly different from this healthy adult. Future researchers should investigate these findings further. The fact that the number of stereotypic items for a healthy adult woman was higher than the number for a healthy adult man may have caused this finding due to a statistical bias. Different results might occur if the numbers of stereotypic items were equal as discussed by Widiger and Settle (1987), who claimed that this difference was a statistical bias, based on the number of socially desirable traits included on the Stereotype Questionnaire.

Our findings also suggest that the gender role stereotypes used by counselors-in-training appear not to have changed much since the early 1970s. Gender role stereotypes remain different for a healthy adult man and a healthy adult woman, as well as different for a healthy adult woman and a healthy adult, sex unspecified. On the other hand, the results are not as clear-cut as those found in the earlier work of Rosenkrantz et al. (1968) and Broverman et al. (1970, 1972). The current results do not suggest that counselors-in-training view a healthy adult man as the "gold standard." Rather, the results suggest that they hold two different standards of mental health—one for healthy women and one for healthy men. These two standards are highly traditional in terms of gender role stereotypes. The most obvious change in these stereotypes is the addition of several competency traits in the description of a healthy adult woman.

General Discussion

To conclude, several changes have occurred in gender role stereotyping over the past 35 plus years. The characterization found for healthy adults was sketchy at best, but included mostly traditionally masculine traits. These traits appear to be the core descriptors of mental health, as three of the five traits ("strong," "independent," and "enjoys a challenge") were included among the stereotypic items for men and women, as well as for an adult, sex unspecified. The other two traits were not shown to overlap between healthy men and healthy women. One, "understanding of others," was shown to be stereotypically feminine, whereas

the other, "knows the way of the world," appeared to be a stereotypically masculine item. It is important that mental health is no longer solely considered synonymous with masculine traits, as one traditionally feminine trait is now viewed as mentally healthy. Additional research focused on healthy adults might help to clarify the current standard for mental health.

Another important finding is the fact that the feminine gender role stereotype has changed, but the masculine gender role stereotype remains consistent with earlier research. Women now are expected to possess both nurturing and competency traits, which suggest that acceptable behaviors for women are broader and more androgynous. In contrast, men's traditional gender role expectations continue to focus solely on competency traits and overlap with most of the characteristics of a healthy adult. But, even though the gender role stereotype for women has changed, a healthy adult woman was still found to be significantly different from both a healthy adult man and a healthy adult, sex unspecified.

Social desirability ratings also have not changed over the last 30 plus years. What are classified as socially desirable traits seem to be deeply, culturally engrained. Given the inflexibility of the social desirability of traits, characteristics, and behaviors, the definition of femininity and masculinity needs to become more androgynous in order to create one standard of mental health. This change would help to eliminate the double standard of mental health for women and men.

Although no differences were found when in comparisons of previous research with between and within group designs (Broverman et al., 1970, 1972; Nesbitt & Penn, 2000; Phillips & Gilroy, 1985; Rosenkrantz et al., 1968; Widiger & Settle, 1987), replication of Study 2 with a between subjects design would help to confirm the results found here. Another methodological issue here is the interrelationships among the many items. Many of these items were endorsed at fairly high levels but did not reach the criterion level we set. Examination of how these traits cluster together might reveal the interrelationships among items.

One additional methodological issue remains, and it is related to the development of items for the questionnaire. Four items on the Stereotype Questionnaire were not effectively changed into unipolar items: "decisive"–"indecisive," "relies on self"–"relies on others," "very careful"–"very careless," and "acts on logic rather than feeling"–"acts on feeling rather than logic." A socially desirable pole was not found for the last item. If two items were created from this one item, such that they were worded as "acts on logic rather than feeling"–"does not act on logic rather than feeling" and "acts on feeling rather than logic"–"does not act on feelings rather than logic" the first item might have been rated as socially desirable. "Decisive" maintained it affiliation with a healthy adult man, but the other two items were not included as a part of the stereotypic scores on any of the Stereotype Questionnaires formats (i.e., healthy adult woman, healthy adult man, or healthy adult). It is possible that, if these items were reworded, differences might appear. That we overlooked these bipolar opposites while overtly attempting to eliminate them demonstrates how deeply ingrained and covert our own stereotypes can become. Future researchers should correctly assign these four items to unipolar formats.

Finally, due to the transparency of the study, participants may have responded in socially desirable ways. Hiding the intent of the study might reveal different results. However, participants still endorsed traditionally socially desirable traits and gender role stereotypes for both women and men. This finding suggests that, although participants may be conscious of gender role stereotypes, they still may hold implicit stereotypes that were revealed in this study. The wording in the directions for the Stereotype Questionnaire may also impact these implicit stereotypes. Participants are asked to think of a "normal" adult (women/man) who is "mature, healthy, (and) socially competent." Here "normal" and "healthy/socially competent" were not separately presented, which might suggest that there is a "typical" adult (woman/man). If one then assumes that there is a typical individual, one might have a tendency to employ stereotypes based on the instructions (Hoffman & Borders, 2001). This confound should be investigated to check whether or not these types of directions can influence the task that was employed in this research.

Future researchers need to investigate whether or not the different mental health standards for women and men are explicitly perceived or only implicitly held by counselors-in-training. Future counselors may be unaware that they hold two different standards of mental health and, therefore, would be expected to deny that they would counsel women differently than men. If future counselors implicitly hold these different standards, course work needs to be focused explicitly on bringing this issue into conscious awareness.

Acknowledgements We thank Mathew Winter for all his help in earlier presentations, data entry and discussions regarding potential outcomes. In addition, we thank Adrian Tomer and the anonymous reviewers for many insightful comments.

References

Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. (1970). Sex-role stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology, 34*, 1–7.

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues, 28*, 59–78.

Crawford, M., & Unger, R. (2000). *Women and gender: A feminist psychology* (3rd ed.) New York: McGraw-Hill.

De Lisi, R., & Soundranayagam, L. (1990). The conceptual structure of sex role stereotyping in college students. *Sex Roles, 23*, 593–611.

Hoffman, R. M., & Borders, L. D. (2001). Twenty-five years after the Bem Sex-Role Inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counseling and Development, 34*, 39–55.

Lerner, R. M. (Ed.). (2002). *Concepts and theories of human development* (3rd ed.). Mahwah, New Jersey: Erlbaum.

Nesbitt, M. N., & Penn, N. E. (2000). Gender stereotypes after thirty years: A replication of Rosenkrantz et al. (1968). *Psychological Reports, 87*, 493–511.

Phillips, R. D., & Gilroy, F. D. (1985). Sex-role stereotypes and clinical judgments of mental health: The Brovermans' findings reexamined. *Sex Roles, 12*, 179–193.

Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., & Broverman, D. M. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology, 32*, 287–295.

Widiger, T. A., & Settle, S. A. (1987). Broverman et al. revisited: An artifactual sex bias. *Journal of Personality and Social Psychology, 53*, 463–469.