Running head: TAGRAS

The Traditional-Antitraditional Gender-Role Attitudes Scale (TAGRAS):

Development and Validation

Ulrich Klocke and Pia Lamberty

Please do not cite literally without permission of the authors. Please cite as follows:

Klocke, U. & Lamberty, P. (2015). The Traditional-Antitraditional Gender-Role Attitudes Scale (TAGRAS): Development and Validation. Manuscript submitted for publication, Institute of Psychology, Humboldt-Universität zu Berlin, Germany.

You are free to use the TAGRAS when you cite it properly. We would be happy to be informed about your results, experiences, and suggestions.

Dr. Ulrich Klocke

Humboldt-Universität zu Berlin

Faculty of Life Sciences

Institute of Psychology

Social and Organizational Psychology

Rudower Chaussee 18

12489 Berlin, Germany

Tel.: +49-30-2093-9335

Fax: +49-30-2093-9332

E-mail: klocke@hu-berlin.de

Pia Lamberty, B.Sc.

Universität zu Köln

Faculty of Human Science

Department of Psychology

Social Psychology: Social Cognition

Richard-Strauss-Str. 2

50931 Köln, Germany

Authors' Note

Abstract

Existing questionnaires usually measure gender role attitudes on a continuum between traditional and egalitarian attitudes. Based on a content analysis of eleven existing questionnaires, we developed the Traditional-Antitraditional Gender Role Attitudes Scale (TAGRAS) to extend this continuum to antitraditional attitudes - a preference for women showing male-typed and men showing female-typed behavior. Two studies with adolescents and one study with adults confirmed that antitraditional attitudes exist. All items loaded on one factor and the TAGRAS has acceptable internal consistency and retest reliability. Convergent validity was demonstrated by showing that men have more traditional attitudes than women and that right-wing political attitudes, social dominance orientation and religious fundamentalism were positively related to traditional attitudes, whereas acceptance of gender-fair language, motivation to act without prejudice, and education level were negatively related to them. Further, the TAGRAS was able to predict explicit and implicit attitudes to lesbians and gays as well as (peer-reported) discriminatory behavior towards lesbians, gays, and gender non-conforming individuals. Implicit attitudes to heterosexuals were not predicted, which is evidence for the TAGRAS's discriminant validity. In the discussion, we suggest future research making use of the fact that the TAGRAS assesses a broader range of gender role attitudes than existing questionnaires.

Keywords

The Traditional-Antitraditional Gender-Role Attitudes Scale (TAGRAS):

Development and Validation

How do people feel about a woman playing soccer or a man who cries when being hurt? Do they prefer a female or a male flight attendant, a female or a male construction worker? As these behaviors or occupations are more typical for one sex than the other, the answers to these questions reflect our attitudes to gender roles. *Gender roles* are societal expectations of people based on their sex, i.e. differential expectations regarding men and women (Eagly, 1987). Thus, *gender role attitudes* are the evaluations of these differential expectations and of target persons who either fulfill or do not fulfill them. People with *traditional* gender role attitudes like or admire men and women who fulfill traditional gender roles more than men and women who act against them, and people with *egalitarian* gender role attitudes do not differentially evaluate men and women in this regard.

But are gender role attitudes only on a continuum between traditional and egalitarian? As gender role attitudes have become more liberal in recent decades (Twenge, 1997, for a meta-analysis of studies in the US), it is likely that some individuals like or admire men and women who act against traditional gender roles, e.g. they prefer men over women to care for children and women over men to become managers. In 2013, 20% of a US sample indicated that they would prefer a female boss if they had the choice (Riffkin, 2014). We call these preferences *antitraditional* gender role attitudes. People with antitraditional gender role attitudes evaluate someone who violates gender norms even more positively than people with egalitarian gender role attitudes. As such, antitraditionalists might be more accepting of lesbians, gays, and transgender people. On the other hand, antitraditionalists might also create a new kind of role stress by disapproving of men who focus on their career instead of their children or women who focus on their children instead of their career. Thus, it seems to be important to not only distinguish between traditional and egalitarian gender role attitudes, but to also identify antitraditional attitudes when asking people about gender roles.

However, existing questionnaires usually do not distinguish between egalitarian and antitraditional gender role attitudes. Thus, they might not cover the full range of gender role attitudes and are prone to ceiling effects (Moradi & Parent, 2013), limiting their capability to predict other attitudes or behaviors. In the present article, we introduce a new scale, the *Traditional-Antitraditional Gender-Role Attitudes Scale (TAGRAS)*. The TAGRAS measures traditional, egalitarian, and antitraditional gender role attitudes. It includes a range of gender-role specific behaviors from different domains, i.e. it is not restricted to the work and home domains. It measures gender role attitudes independently of a general preference for attributes associated with masculinity (e.g. agentic behavior) or femininity (e.g. communal behavior). Its aim is less obvious to respondents, which might reduce the problem of socially desirable answers. Finally, it is suitable for adolescent as well as adult respondents and uses contemporary wording, whereas at least some of the items in existing questionnaires seem outdated. In the following article, we will discuss the existing questionnaires and their strengths and weaknesses and show the relative advantages of the TAGRAS.

Existing Scales on Gender Role Attitudes

Existing scales use different approaches to measure gender role attitudes (for an overview, see Table 1). They differ with regard to whether they use unipolar or bipolar scales and as to whether each item focuses on only one sex, compares both sexes or asks directly about attitudes toward gender egalitarianism.

Scale type *unipolar one sex*: Unipolar rating scales on the appropriateness of attributes for one sex (Pleck, Sonenstein, & Ku, 1994; Thompson & Pleck, 1986) usually refer each attribute to only one sex and do not systematically assess the attribute for both sexes. For example, participants were asked how much they agree with the statement "It is essential for a guy to get respect from others" (Pleck et al., 1994). However, this question was only asked with regard to "guys." It remains unclear whether participants would expect the same behavior for women. High approval of this item might therefore not only measure traditional

gender role attitudes but also a general appreciation of agentic behaviors such as assertiveness and reputation (agentic attitude). In fact, participants might even appreciate women who are more respect-seeking than men, i.e., they might have antitraditional gender role attitudes. Rejecting the item is ambiguous as well: it could mean that the respondent has either egalitarian or antitraditional gender role attitudes.

Scale type *unipolar comparison*: Other instruments directly compare different attributes' appropriateness for men and for women within (at least some of) their items, e.g. "It bothers me more to see a woman who is pushy than a man who is pushy" (Kerr & Holden, 1996; other instruments with the same scale type: King & King, 1997; Krampen, 1979; Larsen & Long, 1988; Spence & Helmreich, 1997). This strategy avoids confounding gender role attitudes with general agentic or communal attitudes. However, it is not possible to distinguish egalitarian from anti-traditional gender role attitudes in this scale type either. Participants who reject the item might be similarly bothered by men and women who are pushy, but they might also be bothered even more by men who are pushy than by women who are pushy. A further problem stemming from a direct comparison of men and women within the same item is that the aim of the measure is obvious. Therefore, participants are more influenced by their perception of relevant others' expectations of which gender role attitude is acceptable (social desirability, Chang, 1999; Jean & Reynolds, 1984).

Scale type *unipolar equality*: Other instruments use at least some of their items to directly ask for participants' attitudes towards gender equality (Kerr & Holden, 1996; King & King, 1997), e.g. "It is just as important to educate daughters as it is to educate sons" (Larsen & Long, 1988). This scale type has the same problems as scale type *unipolar comparison*: participants who reject this item might have traditional gender role attitudes, favoring the education of sons compared to daughters, but they might also have antitraditional gender roles, favoring the education of daughters compared to sons. The same score can represent two different kinds of attitudes, which raises doubts as to the validity of this scale type.

Similarly as with the scale type *unipolar comparison*, the aim of the measure is obvious, which might increase the likelihood of socially desirable answers.

Scale type *bipolar comparison*: The Gender Role Egalitarian Attitudes Test (GREAT, Chang, 1999) uses bipolar scales and asks about the appropriateness of behaviors for women versus for men, e.g. "be a leader" followed by a 9-point scale with four points (1 to 4) meaning "more for men", a middle point (0) meaning "same" and other four points (1 to 4) meaning "more for women." Thus, the GREAT is able to distinguish between traditional, egalitarian and anti-traditional gender role attitudes. Chang (1999) assumes his test to be less influenced by social desirability because "the egalitarian ideology" in his items is "made less explicit" (p. 724). However, in the GREAT, as was also true of the previous two scale types, participants are directly focused on their evaluation of gender differences. Thus, the aim of the instrument is quite obvious. Furthermore, it might take more effort for participants to directly compare expectations of men and women within a single item and be easier to focus on only one sex in each item. Finally, the test is restricted to two life domains, work and home, and assumes that work belongs to the male and home belongs to the female gender role. However, gender roles influence other life domains as well, such as leisure or dating, and also pertain to interpersonal behavior in general (independent of domain). In addition, many jobs belong to the female gender role (e.g. nurse) and some home activities belong to the male gender role (e.g. fixing home appliances).

Overview of the Present Studies

We developed items for the TAGRAS on the basis of a content analysis of eleven existing questionnaires on gender role attitudes and gender role identity. Different from existing scales, we presented each gender-typed behavior twice and asked for separate evaluations when shown by a woman or by a man (see scale type *bipolar separate* in Table 1). We conducted three studies, the first two with adolescents, the third with adults. Participants answered the TAGRAS together with other questionnaires in order to estimate its

correlations with similar (convergent validity) and different constructs (discriminant validity) and to predict attitudes and behavior (predictive validity). Study 1 was used to select the 11 most appropriate item pairs from an initial set of 44 item pairs (88 items). In Studies 2 and 3, the factor structure and internal consistencies were affirmed and complemented by data on stability (retest reliability, Study 2). The TAGRAS's convergent validity was tested by correlations with sex (Studies 1, 2, and 3), education level (Studies 2 and 3), social dominance orientation, religiosity and religious fundamentalism (Study 2), acceptance of gender-fair language, the motivation to act without prejudice and political attitude (Study 3). The TAGRAS's predictive validity was tested by (longitudinally) predicting attitudes and peer-rated behavior towards lesbians and gays controlling for sociodemographic factors (Study 2). The TAGRAS's discriminant validity was tested by correlations with religious reflexivity and attitudes toward heterosexuals (Study 2).

The Traditional-Antitraditional Gender-Role Attitudes Scale (TAGRAS)

Item development

We defined our item population by considering eight existing questionnaires on gender role attitudes (Chang, 1999; Kerr & Holden, 1996; King & King, 1997; Krampen, 1979; Larsen & Long, 1988; Pleck et al., 1994; Spence & Helmreich, 1997; Thompson & Pleck, 1986) and three existing questionnaires on gender role identity, i.e. femininity and masculinity (Bem, 1974; Spence, Helmreich, & Stapp, 1975; Spence, Helmreich, & Holahan, 1979) developed mostly in the United States but also in Beijing and Hong Kong (Chang, 1999) and Germany (Krampen, 1979). This procedure was expected to lead to an initial item pool that covered the male and female gender roles as comprehensively as possible. The 272 items from the 11 questionnaires were categorized with regard to the personal attribute expressed by their content. For example, the items "forceful" (Bem, 1974), "it bothers me more to see a woman who is pushy than a man who is pushy" (Kerr & Holden, 1996), and "in some kinds of situations a man should be ready to use his fists, even if his wife or girlfriend

would object" (Thompson & Pleck, 1986) were categorized under the attribute "aggressive". This categorization resulted in 29 attributes (with 18 miscellaneous and 4 unspecific items), which were further categorized into four dimensions (see Table 2). Out of these categories, a representative set of 44 items was developed by creating at least one item per attribute and more than one item for attributes that were represented by a higher number of items on existing questionnaires. The items describe specific behaviors (e.g. "becomes a manager in a large company"), not personal traits or states (e.g. ambitious), and were always used to describe both "the ideal man" and "the ideal woman," thereby creating item pairs.

The 44 item pairs and their instructions were pretested for understanding with 16 German students (13 to 16 years old), 14 from middle schools (*Realschule*) and two from academic high schools (*Gymnasium*). These students answered the questionnaire and had the opportunity to indicate instructions or items that they found difficult to comprehend. Three of the 16 students were asked in an interview to restate the item content in their own words and to indicate what exactly was difficult to understand. Based on this feedback, some items and the instructions were modified.

Presentation of the Items in the Questionnaire

The TAGRAS presents bipolar scales that focus on only one sex per item, but combine each attribute with men as well as with women in different items (see scale type *bipolar comparison* in Table 1). The items are introduced as follows: "Next, we are interested in your concept of an ideal woman. It is not about a particular woman, but about what behavior you generally find good in women. How do you find it when a woman does the following? Please check the answer which you agree with most." The same paragraph is shown for men as the target persons; in the online version of the test, the order of the two paragraphs is counter-balanced. Participants can evaluate each behavior and sex on 5-point scales, anchored by -2 (*very bad*), -1 (*rather bad*), 0 (*neutral*), +1 (*rather good*), +2 (*very good*). In contrast to existing instruments, the TAGRAS does not ask for a direct comparison

of participants' expectations of men and women within the same item. Thus, the aim of the instrument is less obvious, which should reduce the influence of social desirability.

Estimation of Test Scores

By averaging the differences between the evaluation of men and women for each behavior, the TAGRAS is able to assess gender role attitudes independent of general agentic and communal attitudes. Each difference is calculated by subtracting the score of the gender-atypical behavior (e.g. "she plays soccer") from the score of the corresponding gender-typical behavior (e.g. "he plays soccer"). By interpreting not only the level but also the sign of the average difference score, the TARGAS permits the independent measurement of traditional (positive score), egalitarian (zero score), and anti-traditional (negative score) gender role attitudes.

Translation from German to English

We developed and applied the TAGRAS in the German language. In order to develop an English version, two professional English native translators independently translated the TAGRAS. A single English version was developed and translated back into German by three native German speakers who had lived for several years in English-speaking countries. When translations differed from each other or the back translation differed from the original version, we discussed the differences with the translators and back translators and agreed upon an English version that was most similar to the original with regard to its connotative meaning.

Study 1: Empirical Item Selection

In Study 1, we presented the 44 item pairs to a sample of 77 German pupils in order to select the items that most appropriately measure gender role attitudes using factor and reliability analyses.

Method

      Participants.

      Two hundred and seven German pupils were recruited by snowball sampling via e-mail with the incentive to participate in a lottery for five Amazon or iTunes vouchers (20 € each). Of these, 77 randomly selected pupils completed the questions on gender-role attitudes; the others completed other parts of the questionnaire (see below). Four participants were excluded from the analysis because they had indicated comprehension problems at the end of the questionnaire and/or answered the gender role questions in less time than half of the median of all participants. Two participants were excluded because they had only evaluated one sex. The mean age of the remaining 71 participants was $M = 16.9$ years (range: 14 to 19 years, $SD = 1.6$ years); 41 were female and 30 male; 63 attended an academic high school (*Gymnasium*) and 8 other secondary schools (3 *Integrierte Sekundarschule*, 2 *Berufliches Gymnasium*, 1 *Gesamtschule*, 1 *Realschule*, 1 *Hauptschule*).

      Procedure and Measures.

      The online questionnaire was introduced as a survey on different issues, "such as sexual diversity, prejudice, discrimination or the 'ideal women'". Informed consent was obtained prior to the completion of the questionnaire. Once participants indicated having understood the information on anonymity and their right to withdraw at any time, the questionnaire started. Before participants were asked about their gender role attitudes using the TAGRAS, they indicated their sex, age, school type, class and were asked to report the frequencies of different behaviors from one of their classmates and their class teacher. The items referred to behavior toward lesbian, gay, transgender and gender non-conforming classmates as well as generally hostile and friendly behavior.

Results

First, for each behavior, the scores for the "ideal woman" as a target were subtracted from the scores for "the ideal man" as a target. Thus, a positive score implies that the behavior is preferred (or less disliked) for men and a negative score implies that the behavior is preferred (or less disliked) for women. A principal component analysis was conducted. The scree-test indicated a one-factor solution as most appropriate. Therefore, a one-factor solution was selected, which explained 22% of the variance. Nineteen difference scores had loadings with absolute values higher than .50. All 19 represented attributes from just two of the four dimensions that had been extracted from the 11 existing questionnaires: *strong-weak* and *traditional task allocation* (see Table 2). From these 19 item pairs, 11 were selected that represented 10 different attributes and thus maximally diverse content with regard to gender roles. A second principal component analysis with the 11 selected item pairs was performed (see Table 3), which yielded a clear one-factor solution with the first factor explaining 50% of the variance.

For three of the 11 selected behaviors, participants' average evaluations were in line with traditional gender roles: Specifically, participants much preferred men than women to pay the bill on a date (item pair 3). Furthermore, to cry when something has hurt ones feelings very much was deemed more preferable for women than for men (8), and women were less disliked for becoming a professional hairdresser (1). Interestingly, for item pair 5, participants' average evaluations were in line with antitraditional gender roles because they preferred men to women to take over the cleaning of the apartment. For the remaining seven behaviors, participants in Study 1 held on average egalitarian attitudes, because the evaluations did not significantly differ between men and women as target persons.

The 11 difference scores were averaged in order to get a measure for traditional versus antitraditional gender role attitudes. Before averaging, the scores of item pairs 1, 2, 5, and 8 were reversed (multiplied by -1) because they represent female gender role behaviors,

which means that higher evaluations of men compared to women imply antitraditional gender role attitudes. The scale was normally distributed (Kolmogorov-Smirnov $D = 1.24$, $p = .208$) and internally consistent (Cronbach's $\alpha = .89$). On average, participants had egalitarian gender role attitudes ($M = 0.10$, $SD = 0.82$, 95% CI [-0.10, 0.29]); 31 participants had scores below 0 (antitraditional), 30 had scores above 0 (traditional) and 10 had exactly a score of 0. Boys held more traditional gender role attitudes ($M = 0.42$, $SD = 0.95$) than did girls ($M = -0.14$, $SD = 0.63$), $t(1,69) = 2.97$, $p < .01$. Age did not correlate with gender role attitudes ($r = -.09$, $p = .767$).

Discussion

The distribution of gender role attitudes shows the importance of distinguishing not only between more or less traditional attitudes. Nearly half of our participants had antitraditional gender role attitudes, i.e. they preferred men to show female gender role behavior and women to show male gender role behavior. This finding points to the importance of using a questionnaire that is able to distinguish between traditional, egalitarian and antitraditional attitudes. Also, when focusing on behaviors instead of participants, the results of our questionnaire found the whole range of traditional, egalitarian and antitraditional attitudes: some behaviors were rated in accordance with their traditional gender role, e.g. on average it was preferred for men to pay the bill on a date rather than women. For other behaviors, expectations were reversed, e.g. it was also preferred for men to take over the cleaning of the apartment, which is traditionally a female-typed behavior. But even though some behaviors were evaluated traditionally, some egalitarian, and some antitraditionally on average, in a factor analysis, all selected item pairs loaded high on the same factor measuring individual gender role attitudes.

Initially, we had sampled items from four different dimensions of gender-role specific behavior (see Table 2). However, only items from two of these dimensions loaded high enough to be included in the TAGRAS: *gender-specific task allocation* and *strong vs.*

weak. This finding suggests that these two dimensions are at the core of a traditional vs.

antitraditional gender role attitude whereas the evaluations of being good vs. bad or inhibited

vs. uninhibited are nowadays less related to gender role attitudes.

Study 2: Replication of Factor Structure, Reliability and Validity (Pupils)

For Study 2, we analyzed the items selected in Study 1 using a larger sample of

pupils, which was representative of the different school types in Berlin, Germany. In addition

to factor structure and internal consistency, we assessed TAGRAS's stability over nine

months as a conservative measure of retest reliability. Furthermore, we tested the scale's

convergent validity using correlations with sex and education as well as religiosity, religious

fundamentalism and social dominance orientation as other aspects of conservatism.

These measures were chosen because male participants have been found to hold

more traditional gender role attitudes than female participants (Frieze et al., 2003, for the

United States, Slovenia, and Croatia), and we already confirmed this for the TAGRAS in

Study 1. In addition, participants with a lower education level (Twenge, 1997), higher

religiosity and religious fundamentalism (Aidala, 1985) as well as a higher social dominance

orientation (Russell & Trigg, 2004; B. Whitley & Ægisdóttir, 2000) have more traditional

gender role attitudes. We tested the TAGRAS's discriminant validity via the correlation with

religious reflexivity, which is a more liberal kind of religiosity (Besecke, 2001; Huber, 2008).

Finally, we tested its predictive validity by regressing attitudes and peer-reported behaviors

towards lesbians, gays, and gender non-conforming persons on sociodemographic

characteristics and the TAGRAS. Lesbians and gays describe themselves and their interests as

less gender-role conforming (Lippa, 2005) and are perceived as less gender-typical than

heterosexuals (Stern, West, Jost, & Rule, 2013). Thus, gender role attitudes should predict

attitudes and behaviors towards gays and lesbians and, indeed, people with more traditional

gender role attitudes tend to be more prejudiced against lesbians and gays (meta-analysis of

B. E. Whitley, Jr., 2001).

Method

Participants.

From all regular secondary schools in Berlin, Germany, we randomly sampled 25 schools stratified by district and type of school. The schools' principals received a joint letter (postal and e-mail) by the university and the local school authority asking them to participate in a voluntary study on the acceptance of sexual diversity focusing on ninth or tenth graders. If a school refused to participate, we sampled another school from the same category. This resulted in 51 schools contacted via letter, phone and/or email, and 10 secondary schools with 26 classes and 513 ninth and tenth graders participated. Of these, 32 students were not included in the analyses because 19 reported that they had problems understanding the questionnaire, 14 showed insufficient effort in responding (identified by intra-individual correlations of 15 psychometric antonyms as described by Huang, Curran, Keeney, Poposki, & DeShon, 2011) and 3 showed graphical patterns in their answers. Of the remaining 481 ninth and tenth graders, 45% were female, 49% male and 6% gender unknown; the mean age was $M = 15.2$ years ($SD = 1.0$); 36% were affiliated with a Christian church, 18% with Islam, 3% with other religions, and 43% had no religious affiliation. The distribution among the different school types was comparable to the population distribution in Berlin: 8% secondary modern school (*Hauptschule*, Berlin: 12%), 20% middle school (*Realschule*, Berlin: 17%), 44% academic high school (*Gymnasium*, Berlin: 41%), and 28% comprehensive school (*Gesamtschule*, Berlin: 28%).

Nine months later, 443 of the initial 481 students (25 of 26 classes) participated in a follow-up survey. Of these, 94 were not included in the analyses because 42 reported that they had been distracted by others, 27 reported that others had seen their answers (questions not asked at Time 1), 27 showed insufficient effort in responding, 15 reported that they had had

problems understanding the questionnaire, and 14 showed graphical patterns in their answers, leaving 351 students for a longitudinal analysis.

Procedure.

Prior to data collection, class teachers received information about the survey. Data were collected during school time (first survey: double lesson = 90 min; follow up: single lesson = 45 min). At the first measurement point, the survey was introduced by a member of the research team in 20 classes and by a teacher in six classes. Students were asked to answer a questionnaire about their opinions on different issues, including on gays and lesbians. It was stated that participation was voluntary and that they were free to withdraw whenever they wished. If possible, data were collected online using school computers (311 questionnaires), otherwise by paper and pencil (190 questionnaires). The questionnaire started by asking students about the meaning of different terms related to sexual diversity, e.g. "sexual orientation," and these terms were explained to assure that they were understood in the questionnaire. Afterwards, students' attitudes regarding different sexual identity groups were assessed; the order of the explicit and implicit measures (see below) was counterbalanced. Following this, the TAGRAS was completed together with a number of other measures, which represent possible predictors, and consequences of attitudes towards sexual diversity. In order of assessment, these were: social dominance orientation, gender role attitudes (TAGRAS), religiosity, behavior of two classmates, education level, economic situation, sex, age, and comprehension of the questionnaire.

Measures.

Scale scores were estimated by averaging items for participants who had answered enough items to reach a Cronbach's α of at least .60; otherwise the participant was excluded from the respective analysis.

Gender role attitudes. Gender role attitudes were measured by the 11-item version of the TAGRAS as developed in Study 1. The order of the items was as shown in Table 3.

Sociodemographic variables. We measured economic situation, education level, and age. The students' *economic situation* (Cronbach's $\alpha = .73$) was measured by averaging the z-scores of two subjective and ten objective items (adapted from Gerecht, Steinert, Klieme, & Döbrich, 2007; Kunter et al., 2002). Subjectively, students were asked to estimate their own [their family's] financial situation compared to other students' [other families'] situation in Germany on 5-point scales. Objectively, they were asked how much money they received monthly and can spend as preferred, if they have a room [computer, smart phone, audio player] of their own, if their family owns a garden [house or flat] and how many cars, bathrooms and computers their family owns. The numbers of computers, cars, and bathrooms were calculated relative to the number of members of the household. Due to their higher resolution and factor loading, subjective financial situation, number of computers, cars, and bathrooms received a double weight in the aggregation. *Education level* was measured (adapted from Gerecht et al., 2007; Kunter et al., 2002) by averaging the z-scores of school grades in German, mathematics, and the first foreign language, grade retention, number of books at home, number of musical instruments at home, availability of poetry at home, availability of "classic literature (e.g. Goethe)" at home, and parents' school and vocational education (Cronbach's $\alpha = .81$). School grades were made comparable between different school types by using tables of the Berlin Senate Department for Education. The number of books and musical instruments were calculated relative to the number of members of the household. School grades, number of books and parents' vocational education received double weight in the aggregation.

Social dominance orientation. Students were asked how much they approve of six items of Cohr's and Asbrock's scale (2009), e.g. "It is probably okay that some groups are at

the top of society and others at the bottom" (5-point scales from 0 = *not true* to 4 = *very true*, $M = 1.35$, $SD = 0.82$, Cronbach's $\alpha = .76$).

Religiosity. General *religiosity* was measured using the Religiositäts-Struktur-Test (R-S-T, Huber, 2008) with seven items (Cronbach's $\alpha = .87$), e.g. "How strongly do you believe that god or something divine exists?" (5-point scale from 0 = *not at all* to 4 = *very much*) or "How often do you participate in services, shared prayers or temple rituals?" (6-point scale: *more than once a week, once a week, between one and three times a month, several times a year, more rarely, never*. Because items were answered on scales with varying numbers of points, they were z-transformed before aggregation. Furthermore, the 343 pupils who reported to be at least somewhat religious were asked about two further aspects using the R-S-T: *Religious fundamentalism* was measured by six items (5-point scales from 0 = *not true* to 4 = *very true*, $M = 1.23$, $SD = 1.09$, Cronbach's $\alpha = .89$), e.g. "I am convinced that, with regard to religious issues, particularly my religion is right and other religions tend to be wrong" and *religious reflexivity* by three items (5-point scales from 0 = *never/not at all* to 4 = *very often/very much*, $M = 1.64$, $SD = 0.93$, Cronbach's $\alpha = .77$), e.g. "How important is it for you to consider religious questions from different perspectives?".

Behavior towards lesbians, gays and gender non-conforming persons. The behavioral items were developed through an expert survey with 52 participants and pretested with 99 students ($M = 16.6$ years, $SD = 1.7$ years; 55 female, 43 male and 1 "neither nor") who were asked to report how often they had observed the different behaviors in a classmate. In the main study, students received an index of numbers and names of participating students in their class and were asked to describe the behavior of the two students with numbers directly above theirs. They were requested to focus on behavior they had directly observed within the last 12 months. Usually, they received 4-point scales from 0 (*never*) to 3 (*frequently*). When the behavior was only relevant to a specific situation (e.g. "made fun of a girl who had behaved like a boy"), they received 5-point scales from 0 (*never*) to 4 (*each*

time, transformed into 4-point scales for aggregation), and an additional category indicating that the specific situation had never occurred was provided. Students' *discriminatory behavior towards lesbians, gays and gender non-conforming individuals* ($M = 0.43$, $SD = 0.50$, Cronbach's $\alpha = .83$) was measured by seven items, e.g. "used 'gay' or 'fag' ['lesbian/Dyke'] as abusive words" or "made fun of a boy [girl] who had behaved like a girl [boy]". Students' *supportive behavior towards lesbians and gays* ($M = 0.34$, $SD = 0.47$, Cronbach's $\alpha = .66$) was measured by five items, e.g. "said that s/he thinks gays [lesbians] are okay" or "showed disapproval when a person was teased for [being gay/lesbian]". Because both behavioral variables were highly skewed, they were transformed by an inverse transformation reducing their skewness from 1.43 to 0.77 for *discriminatory behavior* and from 2.59 to 1.55 for *supportive behavior*. The validity of the behavioral assessment was supported by a correlation of $r = .79$ between students' peer reports on discrimination (averaged within classes) and their class teacher's report on the frequency of homophobic insults.

Explicit attitudes toward lesbians and gays. We adapted 17 items from existing questionnaires (Hudson & Ricketts, 1980; Seise, Banse, & Neyer, 2002; Simon, 2008; 5-point scales from -2 to +2, $M = 0.20$, $SD = 0.71$, Cronbach's $\alpha = .90$). The affective aspect was measured by asking students to report their feelings about fictitious contacts with lesbian and gay people from -2 (*very uncomfortable*) to +2 (*very comfortable*), e.g. "You learn that one of your friends is lesbian [gay]". The cognitive aspect was measured by asking students to report how much they approve of seven items on scales from 0 (*not true*) to 4 (*very true*, transformed to scores from -2 to +2 for aggregation), e.g. "Lesbian [gay] couples should be allowed to marry with the same rights as in marriages between a man and woman".

Implicit attitudes towards lesbians, gays and heterosexuals. Implicit attitudes were measured using the affect misattribution procedure (AMP, Payne, Cheng, Govorun, & Stewart, 2005) for the 311 students who completed the online survey. Students received 100 Chinese ideographs as target stimuli (200 ms each) and were asked to judge whether an

ideograph was more or less pleasant than the average ideograph. As primes, 100 ms before

the targets, participants saw black and white photos of romantic couples (100 ms each).

Students were told to disregard the primes and only judge the Chinese ideographs. We used

three prime categories (10 female-female, 10 male-male, and 10 female-male couples, each

used three times) and a grey square as a neutral prime. The three categories were comparable

with regard to action (wedding, kissing on mouth, dancing), age, valence, and ethnicity. Only

items with factor loadings above .40 and the second highest loading at least .10 lower were

included. Judgments faster than 150 ms or slower than 3000 ms were excluded from analyses

as well as data from fourteen students who were familiar with Chinese ideographs. *Implicit*

attitudes towards lesbians and gays were measured by the proportion of "pleasant" judgments

after 21 female-female and 25 male-male prime presentations ($M = .47$, $SD = .25$, Cronbach's

$\alpha = .92$). *Implicit attitudes to heterosexuals* were measured by the proportion of "pleasant"

judgments after 17 female-male prime presentations ($M = .70$, $SD = .24$, Cronbach's $\alpha = .84$).

The measure's validity was supported by an interaction effect between implicit attitudes

towards lesbians and gays and the motivation to appear non-prejudiced towards lesbians and

gays on explicit attitudes to lesbians and gays ($\beta = -.14$, $p = .006$): The higher the students'

motivation to appear non-prejudiced, the less their explicit attitudes could be predicted from

their implicit attitudes.

Results

 Single Behaviors.

 As in Study 1, for each behavior, scores for the "ideal woman" as a target were

subtracted from scores for "the ideal man" as a target. Contrary to Study 1, for all behaviors

but one, participants' average evaluations were in line with traditional gender roles (see Table

4): Even more than in Study 1, participants preferred men rather than women to pay the bill

on a date (item pair 3). Men were also much more preferred over women to join the armed

forces and become professional soldiers (4), play soccer (10), and become car mechanics (7), and women were preferred over men to cry when something has hurt her feelings very much (8). Furthermore, participants preferred men to provide for the family and women to be housekeepers (6). They also preferred men to become managers in large companies (9), whereas they preferred women to interrupt their careers for a year to care for their child (2) and to become hair dressers (1). On average, there were no differences in expectations in terms of acting confidently despite being insecure (11). In contrast to Study 1, participants on average did not show antitraditional gender role attitudes for any of the 11 behaviors.

Factor Structure, Reliability, and Distribution.

A principal component analysis with the 11 difference scores confirmed the one-factor solution with a first factor explaining 32% of the variance. Eight difference scores had loadings with absolute values higher than .50 (see Table 4). Cronbach's α was .78 and not affected by excluding the three difference scores with loadings < .50. Thus, to estimate traditional vs. antitraditional gender role attitudes (TAGRAS), all eleven difference scores were averaged after reversing the scores of female gender role behavior, as in Study 1. Data from fourteen of the 479 participants were excluded because they had more than three missing difference scores. Probably due to the large sample size, the Kolmogorov-Smirnov test showed significant deviation from normality ($D = 0.083$, $p = .003$). However, the inspection of the histogram and the Q-Q plot showed no severe violation of normality and the skewness was only 0.51. The correlation between the two measurement times was .62 ($p < .001$), showing that gender role attitudes were quite stable even after nine months and that our measure's retest reliability was satisfactory. As mentioned above, participants in Study 2 on average had traditional gender role attitudes ($M = 0.84$, $SD = 0.75$, 95% CI [0.77, 0.91]); 13% had scores below 0 (antitraditional), 80% had scores above 0 (traditional) and 7% had exactly a score of 0.

Convergent, Discriminant, and Predictive Validity.

The convergent validity of the TAGRAS was demonstrated via its relation to participants' sex, education level, social dominance orientation, religiosity, and religious fundamentalism: As in Study 1, boys held more traditional gender role attitudes ($M = 1.02$, $SD = 0.76$) than did girls ($M = 0.61$, $SD = 0.65$), $t(1, 431.16) = 6.11$, $p < .001$. Age and economic situation did not correlate with the TAGRAS, but students with a higher education level held less traditional gender role attitudes (see Table 5). Social dominance orientation and religious fundamentalism, which are aspects of conservatism, were positively correlated with the TAGRAS and general religiosity showed a small positive correlation. The TAGRAS did not correlate with religious reflexivity, i.e. a more liberal kind of religiosity, which can be interpreted as an index of discriminant validity.

In order to test the scales' predictive validity, multi-level analyses with students nested within classes were performed using SPSS mixed models and restricted maximum likelihood as the estimation method. Behaviors and attitudes toward lesbians, gays, and heterosexuals as dependent variables were predicted by sociodemographic characteristics and the TAGRAS in level 1 random intercept models (see Table 6). To enable a comparison between the effect estimates, all variables were z-transformed before analyses. The predictive validity of the TAGRAS was demonstrated for discriminatory behavior and both attitude measures. Explicit attitudes toward lesbians and gays were predicted by the TAGRAS even longitudinally for the follow-up survey nine months later after controlling for attitudes at the first measurement point. The null effect on implicit attitudes toward heterosexuals can be considered evidence for the scale's discriminant validity.

Discussion

We were able to confirm the factor structure and internal consistency of the TAGRAS even though the more representative sample of Study 2 had more traditional gender role attitudes overall than the participants of Study 1. We further found that the TAGRAS has

acceptable test-retest reliability over nine months. Furthermore, the TAGRAS's correlations with sex, education level, social dominance orientation, and religious fundamentalism were evidence for its convergent validity, and its non-significant relation to religious reflexivity and implicit attitudes to heterosexuals can be considered evidence for its discriminant validity. Explicit and implicit attitudes and even peer-reported behavior towards lesbians, gays, and gender-nonconforming individuals could be predicted by the TAGRAS, showing its predictive validity. The longitudinal effect of the TAGRAS on explicit attitudes towards lesbians and gays, controlling for attitude at Time 1, suggests that traditional gender role attitudes are causally responsible for prejudice against sexual minority groups.

Study 3: Replication of Factor Structure, Reliability and Validity (Adults)

Because Study 1 and 2 focused on the gender role attitudes of adolescents, it remained open whether the factor structure and reliability were comparable for adults. These issues were addressed in Study 3. In addition, we assessed the construct validity of the TAGRAS by measuring the acceptance of gender-fair language (Nagl-Pietris, 2008), the motivation to act without prejudice (Banse & Gawronski, 2003) and political attitudes. In Germany, there is a controversial debate about the use of gender-fair language (Braun, Oelkers, Rogalski, Bosak, & Sczesny, 2007). Up to this point, there is no universally accepted consensus as to whether it should be used and, if so, which specific form is preferable. Gender-fair language is criticized by many as an ideology despite empirical findings showing the existence of a male bias in the German language. The use of masculine pronouns in a generic sense makes women less visible and promotes the cognitive representation of men compared to women (Koeser, Kuhn, & Sczesny, 2014). People with traditional gender role attitudes should be more permissive with regard to this bias because, traditionally, important societal positions have belonged to the male domain. Thus, we predicted a negative correlation between the TAGRAS and the acceptance of gender-fair language. Furthermore, we predicted that people with a high motivation to act without prejudice would be more

motivated to be unprejudiced towards people not fulfilling traditional gender roles. Thus, motivation to act without prejudice should correlate negatively with the TAGRAS.

Method

Participants.

Our sample consisted of 39 women and 43 men recruited by students at the FernUniversität in Hagen, a university in Germany based on distance education. Participants ranged in age from 16 to 67, with an average age of 35.4 years ($SD = 12.5$). The highest education level for most of the participants was a university degree (38%), followed by the *Abitur* (graduation certificate of an academic high school, 35%) and *Realschulabschluss* (graduation certificate of a middle school, 20%).

Procedure.

Participants were asked to take part in an experiment on creativity in unknown situations. The experiment was conducted at the participants' homes. Participants were informed that their answers were anonymous and that they were allowed to withdraw at any time. They started the computer-based experiment with a Stroop Test (MacLeod, 1991) and then received six photos in random order, three showing mixed-gender groups and three only women. The participants were asked to write a description of each photo in three to five sentences. Half the participants were instructed to use gender-sensitive language and half did not receive this information. They then did the Stroop Test again and completed the TAGRAS, the acceptance of gender-fair language questionnaire, the motivation to act without prejudice scale and some sociodemographic questions. Upon completion, participants were thanked and carefully debriefed.

Measures.

Gender role attitudes. Gender role attitudes were measured using the TAGRAS.

Acceptance of gender-fair language. The acceptance of gender-fair language ($M =$ 3.09, $SD = 0.87$, Cronbach's $\alpha = .84$) was measured on 5-point scales from 1 (*agree not at all*) to 5 (*agree very much*) using the following five items (Nagl-Pietris, 2008): "In my opinion, the topic [gender-fair language] is overly emphasized.", "I don't mind gender-fair language.", "I like that job announcements always address both sexes.", "While writing, I pay attention to the use of gender-fair language." and "I pay attention to the use of gender-fair language in my everyday professional life."

Motivation to act without prejudice. We used the MVV-16 (Banse & Gawronski, 2003) with 5-point scales from 1 (*agree not at all*) to 5 (*agree very much*, $M = 3.71$, $SD = 0.49$, Cronbach's $\alpha = .77$), e.g. "You should not laugh at jokes about minorities" or "You should not be guided by your prejudices".

Political orientation. To determine political orientation we used the method of self-characterization in the left/right spectrum. Participants placed themselves on a scale ranging from 1 (*far left*) to 5 (*far right* $M = 2.50$, $SD = 0.61$).

Education level. We asked about the highest level of education (1 = *no graduation*, 2 = *student*, 3 = *Hauptschulabschluss [graduation certificate of a main school]*, 4 = *Realschulabschluss [graduation certificate of a middle school]*, 5 = *Abitur [graduation certificate of an academic high school]*, 6 = *university degree*).

Results

Single Behaviors.

As in Study 1 and 2, for each behavior, scores for the "ideal woman" as a target were subtracted from scores for "the ideal man" as a target. For three of the 11 selected behaviors, participants' average evaluations were in line with traditional gender roles (see Table 4): Specifically, participants preferred men over women to pay the bill on a date (item pair 3) and to provide for the family while the partner cares for the household (6). Women were preferred

over men to cry when something has hurt her feelings very much (8). As in Study 1, for item pair 5, participants' average evaluations were in line with antitraditional gender roles because they preferred men to women to clean the apartment. For the remaining six behaviors, participants in Study 3 had egalitarian attitudes on average because the evaluations did not significantly differ between men and women as target persons.

Factor Structure, Reliability, and Distribution.

As in Study 1 and 2, a principal component analysis on the 11 difference scores confirmed the one-factor solution. The first factor explained 32% of the variance. Nine difference scores had loadings with absolute values higher than .50 (see Table 4). Cronbach's α was .77. Thus, to estimate traditional vs. antitraditional gender role attitudes (TAGRAS), all eleven difference scores were averaged after reversing the scores of female gender role behavior as in Studies 1 and 2. The scale was normally distributed (Kolmogorov-Smirnov $D =$ .088, $p = .182$). As in Study 1, participants in Study 3 on average held egalitarian gender role attitudes ($M = 0.09$, $SD = 0.59$, 95% $CI = -0.04, 0.22$); 39% had scores below 0 (antitraditional), 52.5% had scores above 0 (traditional) and 8.5% had exactly a score of 0.

Convergent Validity.

The convergent validity of the TAGRAS was supported by its relation to participants' sex, education level, political orientation, motivation to act without prejudice and acceptance of gender-fair language. As expected, men held more traditional gender role attitudes ($M = 0.25$, $SD = 0.50$) than did women ($M = -0.08$, $SD = 0.64$), $t(71.59) = -2.59$, $p = .012$. As in Study 2, age did not correlate with the TAGRAS, but participants with a higher education level held less traditional gender role attitudes (see Table 7). Furthermore, participants with right-leaning or center-leaning political attitudes held more traditional gender role attitudes than people with left-leaning attitudes. Furthermore, the TAGRAS's construct validity was corroborated by a medium to strong negative correlation with

participants' motivation to act without prejudice and a strong correlation with their acceptance of gender-fair language.

Discussion

In Study 3, we were again able to confirm the factorial structure and internal consistency of the TAGRAS. Thus, the TAGRAS is not only applicable for adolescents but also for adults. The correlations with sex and education level were similar to Study 2 and corroborate its convergent validity. The correlations with the motivation to act without prejudice, the acceptance of gender-fair language and political attitudes expand the evidence for the TAGRAS's construct validity.

<div align="center">

General Discussion

</div>

The three studies reported here were designed to develop and test a new scale, the TAGRAS, to measure traditional, egalitarian and antitraditional gender role attitudes simultaneously. Whereas existing questionnaires usually only distinguish between traditional and egalitarian attitudes, we showed that a new kind of attitude has also emerged nowadays, which we call antitraditional. Similarly to people with traditional gender role attitudes, people with antitraditional attitudes distinguish between men and women when they evaluate the appropriateness of different behaviors. However, they prefer men over women to show traditionally female-typed behavior, e.g. cleaning the house, and women over men to show male-typed behavior, e.g. leading a large company. In all three studies, participants were not only divided into traditional or egalitarian, but some held antitraditional gender role attitudes. Even in Study 2, which was comprised of students from a representative sample of Berlin schools, 13% reported antitraditional attitudes. On the other hand, the results also show for which specific behaviors traditional gender roles are still highly influential: the strongest difference in evaluations of men and women in all three studies was that men were still much more expected than women to pay the bill on a date (confirming results of a review by Eaton & Rose, 2011).

The TAGRAS was internally consistent and had acceptable retest reliability as well as factorial, construct and predictive validity. In a variety of tests, the TAGRAS performed as expected, and it demonstrated reasonable psychometric properties for adolescents and adults. Replicating existing research on gender role attitudes (Frieze et al., 2003; Twenge, 1997), men and less educated people had higher scores on the TAGRAS than did women and more educated people. People with traditional (vs. antitraditional) gender role attitudes (measured by the TAGRAS) were also more conservative in other respects. They were more social dominance oriented (Russell & Trigg, 2004; B. Whitley & Ægisdóttir, 2000), fundamentalist in religious issues, politically right-leaning, and less motivated to act without prejudice. The TAGRAS also correlated with other gender-specific variables. The less traditional people's gender role attitudes were, the more they were in favor of gender-fair language. Finally, the TAGRAS has predictive validity, demonstrated by its ability to predict implicit and explicit attitudes as well as discriminatory behavior towards lesbians, gays and gender-nonconforming individuals. The predictive validity even emerged after controlling for sex and education level, and for explicit attitudes it still remained significant longitudinally nine months later controlling for attitudes at Time 1. By contrast, there was no effect on attitudes toward heterosexuals, which can be seen as evidence that the TAGRAS also has discriminant validity.

Because of societal changes, existing questionnaires have a problem with ceiling effects (Moradi & Parent, 2013) in that they do not satisfactorily discriminate at the liberal end of the continuum. Our validation of a measure of antitraditional gender role attitudes is particularly important in this context of social change. To our knowledge, only the Gender Role Egalitarian Attitudes Test (GREAT, Chang, 1999) is able to distinguish between egalitarian and antitraditional gender role attitudes. However, the GREAT is restricted to two domains, work and home, whereas gender roles influence more life domains, such as leisure or dating, and also pertain to interpersonal behavior in general. Moreover, the TAGRAS

captures a broader range of behavioral domains than the GREAT, as it was based on a careful content analysis of eleven existing questionnaires. Furthermore, in many scales, including the GREAT, the goal of the measurement is quite obvious because people are asked to directly compare their expectations of men and women within the same item. The TAGRAS measures attitudes towards specific behaviors separately for men and women. Thus, its aim is less obvious for respondents, which might reduce socially desirable responding.

As the TAGRAS is constructed in a way that produces more attitude variance, it might be better able to predict gender-related phenomena than existing questionnaires. For example, it would be interesting to analyze how partnership satisfaction depends on both partners' gender-role identities and gender-role attitudes or how gender-role attitudes influence the evaluation of an individual person. There may be a difference between abstract beliefs in modern gender roles and reactions when faced with an actual woman behaving in a male-typed or an actual man behaving in a female-typed way (Rudman & Glick, 2010).

In the three studies, we only tested the psychometrics for German samples. Future research should apply the TAGRAS in other countries and languages to test whether the eleven item pairs also load on one factor and are internally consistent elsewhere. Researchers can use the TAGRAS to explore differences in gender role attitudes across different cultures. Past research has shown a cross-cultural consensus about which traits are associated with men and women (Williams & Best, 1990), but we know nothing about antitraditional gender roles across different countries and cultures.

Despite a body of research on biased questionnaire responses due to social desirability, it is still difficult to minimize this problem when measuring explicit attitudes. We consider the TAGRAS to be less vulnerable to social desirable responding because, compared to other questionnaires measuring gender roles attitudes, the measured construct is less obvious to respondents. However, further research ought to explore this on an empirical level (Podsakoff, MacKenzie, & Podsakoff, 2012).

Conclusion

As the first scale of its kind, the TAGRAS represents an advance in the study of gender role attitudes. It shows that nowadays, gender role attitudes are not restricted to the continuum between traditional and egalitarian attitudes, but that some people have antitraditional attitudes as well. We hope that future research using the TAGRAS will examine how antitraditional gender role attitudes develop and how they influence affective, cognitive and behavioral reactions, e.g. towards sexual or gender minorities and in gender-related phenomena such as romantic relationships, work-family interfaces, or leadership.

References

Aidala, A. A. (1985). Social change, gender roles, and new religious movements. *Sociology of Religion, 46*(3), 287-314. doi: 10.2307/3710695

Banse, R., & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Psychometrische Eigenschaften und Validität. [The motivation to act without prejudice scale: psychometric characteristics and validity]. *Diagnostica, 49*, 4-13. doi: 10.1026//0012-1924.49.1.4

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162. doi: 10.1037/h0036215

Besecke, K. (2001). Speaking of meaning in modernity: Reflexive spirituality as a cultural resource. *Sociology of Religion, 62*(3), 365-381. doi: 10.2307/3712355

Braun, F., Oelkers, S., Rogalski, K., Bosak, J., & Sczesny, S. (2007). "Aus Gründen der Verständlichkeit ...": Der Einfluss generisch maskuliner und alternativer Personenbezeichnungen auf die kognitive Verarbeitung von Texten. ["Because of comprehensibility…": The influence of generic masculine and alternative personal designations on cognitive processing of texts]. *Psychologische Rundschau, 58*, 183-189. doi: 10.1026/0033-3042.58.3.183

Chang, L. (1999). Gender Role Egalitarian Attitudes in Beijing, Hong Kong, Florida, and Michigan. *Journal of Cross-Cultural Psychology, 30*, 722. doi: 10.1177/0022022199030006004

Cohrs, J. C., & Asbrock, F. (2009). Right-wing authoritarianism, social dominance orientation and prejudice against threatening and competitive ethnic groups. *European Journal of Social Psychology, 39*, 270-289. doi: 10.1002/ejsp.545

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation.* Hillsdale, NJ: Lawrence Erlbaum.

Eaton, A. A., & Rose, S. (2011). Has Dating Become More Egalitarian? A 35 Year Review Using Sex Roles. *Sex Roles, 64*, 843-862. doi: 10.1007/s11199-011-9957-9

Frieze, I. H., Ferligoj, A., Kogovšek, Tina, Rener, T., Horvat, J., & Šarlija, N. (2003). Gender-role attitudes in university students in the United States, Slovenia and Croatia. *Psychology of Women Quarterly, 27*, 256-261. doi: 10.1111/1471-6402.00105

Gerecht, M., Steinert, B., Klieme, E., & Döbrich, P. (2007). *Skalen zur Schulqualität. Dokumentation der Erhebungsinstrumente. Pädagogische Entwicklungsbilanzen mit Schulen (PEB).* [Scales to measure school quality. Documentation of the measurement instruments. Educational development audits with schools]. (Vol. 17). Frankfurt am Main: Gesellschaft zur Förderung Pädagogischer Forschung, Deutsches Institut für Internationale Pädagogische Forschung.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2011). Detecting and Deterring Insufficient Effort Responding to Surveys. *Journal of Business & Psychology, 16*. doi: 10.1007/s10869-011-9231-8

Huber, S. (2008). Der Religiositäts-Struktur-Test (R-S-T): Systematik und operationale Konstrukte [The Religiositäts-Struktur-Test (R-S-T): Classification and operational constructs]. In W. Gräb & L. Charbonnier (Eds.), *Individualisierung - Spiritualität - Religion: Transformationsprozesse auf dem religiösen Feld in interdisziplinärer Perspektive* (Vol. 1, pp. 137-171). Berlin: Lit Verlag Dr. W. Hopf.

Hudson, W. W., & Ricketts, W. A. (1980). A strategy for the measurement of homophobia. *Journal of Homosexuality, 5*, 357-372. doi: 10.1300/J082v05n04_02

Jean, P. J., & Reynolds, C. R. (1984). Sex and attitude distortion: Ability of females and males to fake liberal and conservative positions regarding changing sex roles. *Sex Roles, 10*, 805-815. doi: 10.1007/bf00287389

Kerr, P. S., & Holden, R. R. (1996). Development of the Gender Role Beliefs Scale (GRBS). *Journal of Social Behavior & Personality, 11*, 3-16.

King, L., & King, D. (1997). Sex-role egalitarianism scale. *Psychology of Women Quarterly, 27*, 71-87. doi: 10.1111/j.1471-6402.1997.tb00101.x

Koeser, S., Kuhn, E. A., & Sczesny, S. (2014). Just Reading? How Gender-Fair Language Triggers Readers' Use of Gender-Fair Forms. *Journal of Language and Social Psychology*. doi: 10.1177/0261927x14561119

Krampen, G. (1979). Eine Skala zur Messung der normativen Geschlechtsrollenorientierung. [A scale to measure the normative gender role orientation]. *Zeitschrift für Soziologie, 8*, 254-266.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., . . . Weiß, M. (2002). PISA 2000: Dokumentation der Erhebungsinstrumente. [PISA 2000: Documentation of the measurements]. *Materialien aus der Bildungsforschung, 72*, 1-410.

Larsen, K. S., & Long, E. (1988). Attitudes toward sex-roles: Traditional or egalitarian? *Sex Roles, 19*, 1-12. doi: 10.1007/bf00292459

Lippa, R. A. (2005). Sexual Orientation and Personality. *Annual Review of Sex Research, 16*, 119-153.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203. doi: 10.1037/0033-2909.109.2.163

Moradi, B., & Parent, M. C. (2013). Assessment of gender-related traits, attitudes, roles, norms, identity, and experiences. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez, K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 2: Testing and assessment in clinical and counseling psychology.* (pp. 467-488). Washington, DC, US: American Psychological Association.

Nagl-Pietris, S. (2008). *Sprache und Geschlecht. Geschlechtergerechte Sprache - Theoretische Entwicklung und praxisrelevante Umsetzung Language and gender. [Gender-fair language - Theoretical development and practice-oriented implementation.].* Umpublished Diploma Thesis. Retrieved from

https://www.vhs.at/fileadmin/uploadsrmc/downloads/Service/FGS_Diplomarbeiten/D
ANagl_8.pdf

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes:
Affect misattribution as implicit measurement. *Journal of Personality and Social
Psychology, 89*, 277-293. doi: 10.1037/0022-3514.89.3.277

Pleck, J. H., Sonenstein, F. L., & Ku, L. C. (1994). Attitudes toward male roles among
adolescent males: A discriminant validity analysis. *Sex Roles, 30*, 481-501. doi:
10.1007/bf01420798

Podsakoff, P., MacKenzie, S., & Podsakoff, N. (2012). Sources of Method Bias in Social
Science Research and Recommendations on How to Control It. *Annual Review of
Psychology, 63*(1), 539-569. doi: doi:10.1146/annurev-psych-120710-100452

Riffkin, R. (2014). Americans Still Prefer a Male Boss to a Female Boss: Women are more
likely than men to prefer a female boss. *GALLUP ECONOMY*.
http://www.gallup.com/poll/178484/americans-prefer-male-boss-female-boss.aspx

Rudman, L. A., & Glick, P. (2010). *The Social Psychology of Gender. How Power and
Intimacy shape Gender Relations*. New York: The Guilford Press.

Russell, B., & Trigg, K. (2004). Tolerance of Sexual Harassment: An Examination of Gender
Differences, Ambivalent Sexism, Social Dominance, and Gender Roles. *Sex Roles, 50*,
565-573. doi: 10.1023/B:SERS.0000023075.32252.fd

Seise, J., Banse, R., & Neyer, F. J. (2002). Individuelle Unterschiede in impliziten und
expliziten Einstellungen zur Homosexualität: Eine empirische Studie. [Individual
differences in implicit and explicit attitudes towards homosexuality: An empirical
study]. *Zeitschrift für Sexualforschung, 15*, 21-42. doi: 10.1055/s-2002-25178

Simon, B. (2008). Einstellungen zur Homosexualität: Ausprägungen und psychologische
Korrelate bei Jugendlichen ohne und mit Migrationshintergrund (ehemalige UdSSR
und Türkei). [Attitudes towards Homosexuality: Manifestations and psychological
correlates among adolescence with and without migration background (former USSR

and turkey)]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie,*
40, 87-99. doi: 10.1026/0049-8637.40.2.87

Spence, J. T., & Helmreich, E. D. (1997). The Attitudes toward women scale and attitude
change in college students. *Psychology of Women Quarterly, 21*, 17.

Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role
attributes and their relation to self-esteem and conceptions of masculinity and
femininity. *Journal of Personality and Social Psychology, 32*, 29-39. doi:
10.1037/h0076857

Spence, J. T., Helmreich, R. L., & Holahan, C. K. (1979). Negative and positive components
of psychological masculinity and femininity and their relationships to self-reports of
neurotic and acting out behaviors. *Journal of Personality and Social Psychology, 37,*
1673-1682. doi: 10.1037/0022-3514.37.10.1673

Stern, C., West, T. V., Jost, J. T., & Rule, N. O. (2013). The politics of gaydar: Ideological
differences in the use of gendered cues in categorizing sexual orientation. *Journal of*
Personality and Social Psychology, 104, 520-541. doi: 10.1037/a0031187

Thompson, E. H., & Pleck, J. H. (1986). The structure of male role norms. *American*
Behavioral Scientist, 29, 531-543. doi: 10.1177/000276486029005003

Twenge, J. M. (1997). Attitudes toward women, 1970–1995: A meta-analysis. *Psychology of*
Women Quarterly, 21, 35-51. doi: 10.1111/j.1471-6402.1997.tb00099.x

Whitley, B., & Ægisdóttir, S. (2000). The Gender Belief System, Authoritarianism, Social
Dominance Orientation, and Heterosexuals' Attitudes Toward Lesbians and Gay Men.
Sex Roles, 42, 947-967. doi: 10.1023/A:1007026016001

Whitley, B. E., Jr. (2001). Gender-role variables and attitudes toward homosexuality. *Sex*
Roles, 45, 691-721. doi: 10.1023/A:1015640318045

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multination study (rev.*
ed.). Thousand Oaks, CA, US: Sage Publications, Inc.

Tables

Table 1

Taxonomy of Existing Scales to Measure Gender Role Attitudes

Scale type	Description	Also measures anti-traditional gender role attitudes	Independent from general agentic and communal attitudes	Influenced by social desirability
Unipolar one sex	Unipolar rating scales on appropriateness for one sex	no	no	low-medium
Unipolar comparison	Unipolar rating scales on appropriateness for one sex compared to the other	no	yes	high
Unipolar equality	Unipolar rating scales on attitudes to gender equality	no	yes	high
Bipolar comparison	Bipolar rating scale of appropriateness for for one sex compared to the other	yes	yes	medium
Bipolar separate	Bipolar rating scales of appropriateness for men and for women in separate items	yes	yes	low-medium

Table 2

Categorization of Existing Items Measuring Gender Role Attitudes: Attributes and Dimensions

Dimensions	Attributes
strong-weak	dependent vs. independent, strong vs. weak, patriarchical, persistent, dominant, being a cavalier, determined, leading, confident, athletic, competitive
traditional task allocation	successful in one's career, domestic work, child rearing, male jobs, female jobs, educated
good-bad	aggressive, helpful vs. selfish, empathetic, warm, kind vs. rude, arrogant
inhibited-uninhibited	neat, direct, extraverted vs. introverted, sexualized vs. chaste, risk-taking, emotional

Table 3

Item Analysis (Study 1)

No.	Dimens.	Attribute	German original items	English translation of items	Gen-der role	Differences: Eval. men - eval. women			Eval. men	Eval. women
						Loading	*M*	*SD*	*M*	*M*
1	trad. task allocation	female jobs	wird von Beruf Friseur/in.[a]	becomes a professional hairdresser.	♀	-.58	-0.25 *	0.84	-0.29	-0.03
2	trad. task allocation	child rearing	unterbricht für ein Jahr seine/ihre Berufs-tätigkeit, um für sein/ihr Kind zu sorgen.	interrupts his/her career for a year to care for his/her child.	♀	-.70	-0.01	0.93	1.07	1.10
3	strong-weak	chivalrous	bezahlt bei einem Date die gemeinsame Rechnung.	pays the bill on a date.	♂	.59	1.24 ***	1.41	1.15	-0.12
4	trad. task allocation	male jobs	geht zur Bundeswehr und wird Berufssoldat/in.	joins the armed forces and becomes a professional soldier.	♂	.72	0.09	1.01	-0.75	-0.81
5	trad. task allocation	domestic work	übernimmt zu Hause das Putzen der Wohnung.	At home, s/he takes over the cleaning of the apartment.	♀	-.82	0.59 ***	1.42	0.76	0.17
6	trad. task allocation	successful in one's career	versorgt die Familie, während der/die Partner/in[b] für den Haushalt sorgt.	provides for the family while the partner cares for the household.	♂	.82	-0.24	1.54	0.01	0.21
7	trad. task allocation	male jobs	wird Automechaniker/in.	becomes a car mechanic.	♂	.59	-0.17	1.15	0.04	0.19
8	strong-weak	strong vs. weak	weint, wenn ihn/sie etwas sehr verletzt hat.	cries when something has hurt her feelings very much.	♀	-.72	-0.61 ***	1.25	0.50	1.11
9	strong-weak	leading	wird Manager/in in einem großen Wirtschaftsunternehmen.	becomes a manager in a large company.	♂	.73	-0.07	1.07	0.81	0.83
10	strong-weak	athletic	spielt Fußball .	plays soccer.	♂	.73	0.14	1.09	0.61	0.47
11	strong-weak	confident	gibt sich selbstbewusst, auch wenn er/sie verunsichert ist.	pretends to be self-confident even if she is insecure.	♂	.71	-0.21	1.09	0.35	0.56

Note. The asterisks indicate significant differences between the evaluations of men and evaluations of women.

[a] In the German version, the nouns with a slash (e.g. "Friseur/in") were exchanged by the female form (e.g. "Friseurin") when asking for expectations of women and the male form (e.g. "Friseur") when asking for expectations of men (for an exception, see note [b] in item 6).

[b] In the German version, the male form ("der Partner") was used when asking for expectations of women and the female form ("die Partnerin") when asking for expectations of men.

* $p < .05$. *** $p < .001$.

Table 4

Item Analysis (Studies 2 and 3)

No.	English translation of items	Gender role	Study 2 Differences: Eval. men - eval. women			Eval. men	Eval. women	Study 3 Differences: Eval. men - eval. women			Eval. men	Eval. women
			Loading	M	SD	M	M	Loading	M	SD	M	M
1	becomes a professional hairdresser.	♀	-.46	-0.46 ***	1.19	-0.12	0.33	-.33	-0.06	0.78	-0.54	-0.48
2	interrupts his/her career for a year to care for his/her child.	♀	-.52	-0.62 ***	1.22	0.55	1.15	-.50	-0.02	1.04	1.09	1.11
3	pays the bill on a date.	♂	.36	2.20 ***	1.64	1.42	-0.79	.57	0.68 ***	1.51	0.88	0.20
4	joins the armed forces and becomes a professional soldier.	♂	.74	1.17 ***	1.38	0.66	-0.52	.78	0.18	0.88	-0.66	-0.84
5	At home, s/he takes over the cleaning of the apartment.	♀	-.65	-0.39 ***	1.72	0.05	0.44	-.51	0.45 **	1.38	0.32	-0.13
6	provides for the family while the partner cares for the household.	♂	.52	0.75 ***	1.55	0.86	0.10	.52	0.33 *	1.14	0.46	0.13
7	becomes a car mechanic.	♂	.71	1.11 ***	1.28	0.72	-0.39	.68	-0.02	0.97	0.12	0.15
8	cries when something has hurt her feelings very much.	♀	-.52	-1.05 ***	1.25	-0.09	0.96	-.51	-0.23 *	0.99	0.43	0.66
9	becomes a manager in a large company.	♂	.50	0.29 ***	1.11	1.38	1.08	.62	0.00	1.03	0.59	0.59
10	plays soccer.	♂	.69	1.12 ***	1.41	1.18	0.04	.66	-0.04	1.08	0.49	0.52
11	pretends to be self-confident even if she is insecure.	♂	.42	0.06	0.93	0.92	0.86	.45	0.04	0.76	0.55	0.51

Note. The asterisks indicate significant differences between the evaluations of men and evaluations of women.
* *p* < .05. ** *p* < .01. *** *p* < .001.

Table 5

Construct Validity: Pearson Correlations (Study 2)

	1	2	3	4	5	6	7	8
1. TAGRAS								
2. Male (vs. female) sex	.28 ***							
3. Age	.06	.04						
4. Educational level	-.32 ***	-.08 #	-.38 ***					
5. Economic situation	-.04	.09 #	-.02	.23 ***				
6. Social dominance orientation	.38 ***	.31 ***	.03	-.18 ***	.18 ***			
7. Religiosity	.13 **	-.15 **	.01	-.10 *	-.18 ***	-.04		
8. Religious fundamentalism	.27 ***	-.01	.09	-.24 ***	-.20 ***	.17 **	.73 ***	
9. Religious reflexivity	-.05	-.17 **	.01	.12 *	-.08	-.18 **	.62 ***	.42

Note. * *p* < .05. ** *p* < .01. *** *p* < .001.

Table 6

Predictive Validity: Multi-level Analyses with Z-Standardized Predictors (Study 2)

	Discriminatory behavior towards lesbians, gays, and gender-nonconf. indiv.	Supportive behavior towards lesbians and gays	Explicit attitudes towards lesbians and gays	Implicit attitudes towards lesbians and gays	Implicit attitudes towards heterosexuals
Female (vs. male) sex	-.21 ***	.13 *	.20 ***	-.09	.00
Economic situation	.08	.00	.07	-.01	.02
Education level	.08	.16 **	.08	.12 #	-.11
TAGRAS	.14 *	-.02	-.33 *** [a]	-.28 ***	-.05

Note. [a] Effect is also significant ($\beta = -.11$, $p < .05$) for a prediction of attitudes at Time 2 by predictors at Time 1, controlling for attitudes at Time 1.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Construct Validity: Pearson Correlations (Study 3)

	1	2	3	4	5	6
1. TAGRAS						
2. Male (vs. female) sex	.28 *					
3. Age	-.10	.01				
4. Educational level	-.25 *	-.01	.05			
5. Right (vs. left) wing political attitudes	.39 **	.06	.03	-.34 **		
6. Acceptance of gender-fair language	-.49 **	-.29 **	.21	.38 **	-.34 **	
7. Motivation to act without prejudice	-.41 **	-.26 *	.21	.21	-.20	.29 **

Note. * *p* < .05. ** *p* < .01.