



Open Access Repository

[www.ssoar.info](http://www.ssoar.info)

## Combining Quantitative Experimental Data with Web Probing: The Case of Individual Solutions for the Division of Labor Between Both Genders

Braun, Michael; Meitinger, Katharina; Behr, Dorothee

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

### Empfohlene Zitierung / Suggested Citation:

Braun, M., Meitinger, K., & Behr, D. (2020). Combining Quantitative Experimental Data with Web Probing: The Case of Individual Solutions for the Division of Labor Between Both Genders. *Methods, data, analyses : a journal for quantitative methods and survey methodology (mda)*, 14(2), 1-25. <https://doi.org/10.12758/mda.2020.03>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

# Combining Quantitative Experimental Data with Web Probing: The Case of Individual Solutions for the Division of Labor Between Both Genders

*Michael Braun*<sup>1</sup>, *Katharina Meitinger*<sup>2</sup> & *Dorothee Behr*<sup>1</sup>

<sup>1</sup> *GESIS – Leibniz Institute for the Social Sciences, Germany*

<sup>2</sup> *Utrecht University, Netherlands*

## Abstract

In 2012, a new question was introduced into the International Social Survey Program (ISSP). It asks respondents to indicate what they consider the best division of labor between men and women. In this paper, we propose to assess the validity and cross-national comparability of this new ISSP question, using a mixed-methods approach that combines quantitative experimental data with qualitative probing data. We implemented our experiment in non-probability online surveys in five countries, in which half of the respondents received the original ISSP question and the other half a variant with an additional category saying “Each family should find the solution which works best for them.” In addition, the understanding of “individual solutions” was probed. We report on the understanding of this category.

**Keywords:** Web probing, cross-cultural research, mixed methods, gender roles, ISSP



© The Author(s) 2020. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Since 1985, the International Social Survey Program (ISSP) has conducted studies on different areas of social science research and thereby produced a huge data base for comparisons across countries and time. The majority of questions and items are held constant and kept unchanged over the different replications, but some questions are replaced in order to improve measurement quality or capture new trends. One of the topical modules of the ISSP is on “Family and Changing Gender Roles”, which was fielded in 1988, 1994, 2002, and 2012. The way gender ideology was measured in the earlier surveys has often been criticized for having a traditional slant, focusing exclusively on women and employment, or for having methodological problems (Braun, 2008; Edlund & Öun, 2016). Though, from early on, there have also been attempts by researchers to construct more differentiated instruments that partly also capture subtle sexism (Brogan & Kutner, 1976; Glick & Fiske, 1997; King & King, 1997; Swim et al., 1995), these measure have not been adopted by large scale-comparative surveys.

In order to improve the measurement in the ISSP, in the 2012 round, a new measure for gender ideology was included to address respondents’ preferences for the division of labor between men and women when there are children at home (ISSP Research Group, 2016; Scholz et al., 2014). Six types of preferences were presented as response categories, ranging from *the mother stays at home and the father works full-time* to the opposite division of labor (see further down for more details). Respondents should indicate what, according to their opinion, was the best way to organize the division of labor for a couple. This question forced respondents to single out one specific division of labor between men and women. Such a choice could be difficult for respondents who think that the best solution should be made dependent on additional considerations. For example, some respondents might think that the best solution should depend on the preferences of the partners, their abilities or their earning potential. Such respondents might struggle to choose one of the categories offered to them, and this might encourage superficial and stereotypical answer behavior. Therefore, when designing the items, the ISSP drafting group discussed whether an additional “individual solutions” category should be added. If so, this would give respondents who do not find their preferences represented in the answer categories an appropriate way out without having to either opt for “don’t know” or select one of the substantive categories they do not really approve. The addition of such an “individual solutions” category, however, was eventually declined on the basis of concerns that in particular traditional respondents might use this category in order to avoid an overt disclosure of their traditional stance due to social-desirability considerations.

---

*Direct correspondence to*

Michael Braun, GESIS – Leibniz Institute for the Social Sciences, Germany  
E-mail: michael.braun@gesis.org

As to the new item in the ISSP, this should be thoroughly assessed and checked for measurement equivalence across countries before it is used in substantive research. The most commonly used statistical technique for assessing measurement equivalence is multiple-group confirmatory factor analysis (MGCFA, Jöreskog, 1971). Latent-class analysis also has a long tradition in this field (Clogg, 1984; for an application to gender-role items of the World Value Survey and the European Values Study, see Knight & Brinton, 2017). Other techniques include correspondence analysis (for an application to the ISSP gender role items, see Blasius & Thiessen, 2006). All these quantitative methods are helpful in deciding whether measures are equivalent across countries but they usually do not allow getting at the causes of non-equivalence.<sup>1</sup> Much can be gained from getting at the causes of non-equivalence as well as from understanding the interpretations of respondents from different countries. Such interpretation patterns can be used in substantive research to avoid wrong conclusions. In addition, these quantitative methods cannot be applied to single items but to multiple-item measures only. Thus, they cannot be used to assess the new ISSP item on the division of labor between men and women. This is where qualitative approaches can and should come in.

Qualitative approaches, in particular cognitive interviews, are helpful to investigate problems in the response process (Beatty & Willis, 2007; Willis, 2005). A variety of probing techniques exist that are used during cognitive interviewing. For example, category-selection probes help to reveal the reasons for the selection of the responses to closed questions (“Please explain why you selected ‘strongly agree’”). Unfortunately, international comparative cognitive studies drastically increase the coordination effort and are quite time-consuming (Willis, 2015) and, thus, are not implemented frequently in research (for exceptions see: Benítez et al., 2018; Fitzgerald et al., 2011; Miller et al., 2011; Thrasher et al., 2011; for a review see Willis, 2015).

However, the conduct of additional web-based studies to capture cross-cultural qualitative information is a potential source of information. “Web probing, that is, the implementation of probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey items” (Behr et al., 2017), is a method to complement quantitative techniques to establish measurement equivalence of items in cross-cultural research (Behr et al., 2017; Meitinger, 2017). In contrast to quantitative approaches that usually presuppose multiple-item measures, cognitive interviewing and web probing can also assess the cross-national comparability of single questions or items. In web probing studies, probing questions can be included in a regular web questionnaire. Behr & Braun (2015), for example,

---

1 While some quantitative approaches, such as multilevel structural equation modeling (MLSEM), can explain noninvariance by introducing macro-level variables in a multilevel analysis (Davidov et al., 2012), they are very demanding (e.g. samples should exceed 50 countries, see Meuleman & Billiet, 2009).

use a “category-selection” probe for a single item on satisfaction with democracy in order to find out which dimensions of democracy this question measures. The authors found that policy outcomes, governance, and aspects of the concrete political system play an important role in all countries of their study and, thus, answers can meaningfully be compared across countries.

Therefore, for assessing the consequences of including vs. excluding an “individual solutions” category, a mixed-methods approach seems to be particularly helpful (Creswell, 2014; Luyt, 2012; van de Vijver & Chasiotis, 2010). In the present case, we propose to combine the analysis of the quantitative survey data of the ISSP with a separate web study in which a split-half experiment with varying response categories was combined with a qualitative component. While the question experiment can inform the decision as to whether an “individual solutions” category matters in principal, the comparison of the web survey data with the data collected as part of the ISSP survey allows answering the question whether our results can be used to draw conclusions for the ISSP survey and its questionnaire.

## Data and Methods

### Sample

We implemented an experiment in non-probability online surveys in Germany, Great Britain, the United States, Mexico, and Spain with a total of 2,689 respondents. Survey participation was restricted to citizens of the respective countries aged 18 to 65. A net sample of approx. 500 respondents in each country was targeted using quotas for age (18-30, 31-50, and 51-65), gender, and education (lower vs. higher education). The panel providers were Respondi ([www.respondi.com](http://www.respondi.com)) and its partners in the respective countries. We met all quotas (see Table A1 in the Appendix for respective quota fields). Data collection was in June 2014. As these are quota samples, standardized response rates cannot be computed (Baker et al., 2010).

The selection of the five countries for the study was motivated by the expectation that in the liberal regime type (here represented by Great Britain and the United States) individuals or institutions outside of the family should not interfere with decisions regarding the roles of men and women in a family (compared to the conservative regime type here represented by Germany, Mexico, and Spain). These expectations should run in parallel to the lower involvement that the state has with regard to families (including the provision of a supporting infrastructure) in the first group of countries. Mexico was included alongside Germany and Spain as a strongly conservative country in which the family itself has a particularly high

importance in providing a support structure that might become relevant when it comes to the division of labor between both genders.

## Questionnaire

The International Social Survey Program (ISSP, ISSP Research Group 2016) asked the following new question in its 2012 “Family and Changing Gender Roles” module to capture respondents’ views on the preferred division of labor between mother and father:

“Consider a family with a child under school age. What, in your opinion, is the best way for them to organize their family and work life?”

- 1 The mother stays at home and the father works full-time.
- 2 The mother works part-time and the father works full-time.
- 3 Both the mother and the father work full-time.
- 4 Both the mother and the father work part-time.
- 5 The father works part-time and the mother works full-time.
- 6 The father stays at home and the mother works full-time.”

In our web survey, half of the respondents received the original ISSP question (see Figure 1), the other half of the respondents received a variant (developed for this experiment) in which an additional category “Each family should find the solution which works best for them” was added. The respondents who selected the additional answer category also received a probing question regarding the reasons for opting for “individual solutions” (see Figure 2).

Thus, the experimental design combines quantitative insights from the split-ballot experiment with qualitative insights from web probing. To ensure the comparability of the probes themselves, we applied the team-driven TRAPD approach for the translation of the probes (Harkness, 2003).

**Consider a family with a child under school age.**

**What, in your opinion, is the best way for them to organise their family and work life?**

- The mother stays at home and the father works full-time
- The mother works part-time and the father works full-time
- Both the mother and the father work full-time
- Both the mother and the father work part-time
- The father works part-time and the mother works full-time
- The father stays at home and the mother works full-time

---

*can't choose*

*Figure 1* Experimental condition without response category “individual solutions”

**Consider a family with a child under school age.**

**What, in your opinion, is the best way for them to organise their family and work life?**

- The mother stays at home and the father works full-time
- The mother works part-time and the father works full-time
- Both the mother and the father work full-time
- Both the mother and the father work part-time
- The father works part-time and the mother works full-time
- The father stays at home and the mother works full-time
- Each family should find the solution which works best for them

---

*can't choose*

**Please explain why you selected "Each family should find the solution which works best for them".**

The question was: "Consider a family with a child under school age. What, in your opinion, is the best way for them to organise their family and work life?"

- The mother stays at home and the father works full-time
- The mother works part-time and the father works full-time
- Both the mother and the father work full-time
- Both the mother and the father work part-time
- The father works part-time and the mother works full-time
- The father stays at home and the mother works full-time
- Each family should find the solution which works best for them
- *can't choose*

*Figure 2* Experimental condition with response category “individual solutions” and category-selection probe

## Translation of Open-ended Answers, Development of the Coding Scheme, and Coding

The Mexican and Spanish answers to the probe were translated into German by professional translators who had been briefed on the particularities of these texts as well as on translation and coding needs (Behr, 2015). The German and English answers were not translated but immediately coded by members of the project team (German native speakers with high proficiency in English).

An elaborated category scheme was developed, which represents the main criteria for the division of labor. This scheme was based on theory and also on the content of the probe responses.

Several theoretical perspectives can be found in the literature and based on these we developed hypotheses informing our probe scheme development. First, we wanted to investigate whether some of the approaches traditionally used to explain the actual household division of labor are also reflected in the reasoning of the respondents. These approaches are the *time-availability approach* that stipulates that spouses who spend more time working outside of the household show reduced participation with housework (Bianchi et al., 2000; Kalleberg & Rosenfeld, 1990) and the *resource-dependency approach* (Bittman et al., 2003; Brines, 1994) which recurs on the bargaining power of the spouses (based e.g. on their income or education) and its use to avoid unwanted housework. Second, we expected respondents to refer to individual preferences and capabilities, that is, what spouses want to do and where they are good at. Third, we surmised that several respondents would not recommend specific role distributions because they think that such decisions are the responsibility of the respective families or depend on the family's financial situation or on how child care can be organized (e.g. the presence of one parent or relatives at home or other alternative childcare arrangements).

The category scheme will be presented further below together with the results. Multiple coding was possible for all categories except for the categories *no generalization possible*, the *substantive rest category*, and *probe nonresponse*.

After the establishment of the final coding by members of the research team, a research assistant not involved in the development and implementation of the coding scheme coded 90% of the probe answers of all countries (while the other 10% were used for training purposes). Inter-rater agreement (between the final coding by members of the research team on the one hand and the research assistant on the other hand) ranged from 96% in Spain to 100% in the United States and Mexico. The high reliability value is likely to be a consequence of the relatively simple coding scheme, both as far as the number and the definitional clarity of the categories is concerned. This means that in more than 9 out of 10 cases, the raters coded a probing answer identically. All discrepancies of coding were discussed in the research team, which then arrived at a final version used in this paper.



## Analytical Strategy

In the following, we first compare the response pattern found in the ISSP data with the pattern revealed by our web survey to assess the general usefulness of our web survey data. Second, we compare the two experimental conditions implemented in the web survey. The first experimental condition asks the question on the best division of labor between father and mother exactly as it was in the ISSP, and the second experimental condition adds the answer category “Each family should find the solution which works best for them”. Third, we report the responses to the category-selection probe regarding which “individual solutions” respondents had in mind when answering the closed question.

## Results

### Replication of the Pattern in the ISSP Data

A comparison of the first split of our web survey (which exactly replicates the original ISSP question) with the ISSP data<sup>2</sup> reveals that the general response pattern is replicated in our web survey (see Tables A2 and A3 in the Appendix). The ISSP and the web samples share the nearly complete lack of support for a role reversal and similar percentages of respondents who opt for the “don’t know” category. In all five countries, the overwhelming majority supports the strict (only the father goes out to work) or moderate variant (the mother has only a complementary work role) of the male-breadwinner model if they are forced to choose among the models presented.

However, the respondents of the web survey seem to be less traditional than the respondents in the ISSP, despite of the quotas we have implemented for age, gender, and education.

Nevertheless, because of the experimental approach taken here, we are confident that the results found on the basis of the probing study can shed light on the ISSP data.

### “Individual Solutions” for the Division of Labor Between Both Genders

Table 1 shows the response distribution of the closed item in the web survey (the preferred division of labor between men and women), where the second split of the web survey contains the additional answer category “Each family should find

---

2 We did not restrict the ISSP data to the age range of the web survey.

*Table 1* Preferred division of labor dependent on the presence of an “individual solutions” category in the different countries (in percent)

	Germany		Great Britain		United States		Mexico		Spain	
	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2
Mother at home, father full-time	20	11	27	14	27	12	22	16	4	3
Mother part-time, father full-time	35	22	28	10	23	11	50	29	21	10
Both full-time	10	4	10	5	23	10	11	8	22	6
Both part-time	23	6	13	4	6	2	15	4	45	15
Father part-time or at home, mother full-time	1	0	1	0	1	1	0	1	0	0
Individual solutions	-	55	-	65	-	61	-	40	-	66
Don't know	13	1	20	2	20	3	2	1	8	0
<i>N</i>	275	264	281	253	266	274	253	292	268	263

*Data source:* Web survey; split 1: original ISSP version, split 2: “individual solutions” category added; original categories “father part-time and mother full-time” and “father at home and mother full-time” collapsed.

the solution which works best for them.” When this individual-solutions category is introduced, clearly more than half of the respondents choose this category, with the only exception of Mexico (40%). In addition, when this answer category is provided, the prevalence of “don’t know” responses drops drastically (from 2-20% to 0-3%). All other divisions of labor are chosen considerably less in the second split (with “individual solutions”) compared to the first split (without “individual solutions”). However, the relative decrease is most marked for the “both part-time” category in most of the countries.

Though our experiment represents a between- instead of a within-subjects design, it seems nevertheless fair to conclude that the individual solutions category contains those respondents who would opt for the “don’t know” option when the “individual solution” option is not available. In addition, the individual solutions category draws from all substantive categories and, in particular, from the “both part-time” response category. Part of the respondents choosing this category seem to use it as a compromise since none of the categories offered match their real preferences. Thus, it is this – not particularly traditional – category which loses support once the individual solutions category is added and not the more traditional answer

categories, as feared among questionnaire developers when designing the new ISSP question.

This can be seen even clearer from Table A4 in the Appendix which shows the web survey results if only the substantive ISSP categories (that is, without “don’t know” and “individual solutions”) are included in calculating the percentages. In all countries but the United States, it is the most traditional answer category that gains relative importance if an “individual solution” category is added (in the United States, it simply makes no difference). This applies to both genders (see Tables A5 and A6 in the Appendix). Table A7 in the Appendix shows the popularity of the individual solutions category in different social groups, in addition to gender. In most countries, those who opt for the individual solutions category are older than those who do not. Those who are married are less in favor of individual solutions than unmarried respondents. However, there are no consistent relationships between the choice of the individual solutions category and respondents’ employment status and their partners’ employment status and whether they have children or not.

Therefore, it is fair to conclude that adding an individual solutions category is not mainly used as an easy escape by traditional respondents who do not want to disclose their position in an overt manner. It might rather be used by those respondents who think that it impossible to opt for only one of the presented divisions of labor, unless more details on the specific situation of the respective family are taken into account.

## **“Individual Solutions” Respondents have in Mind**

What, then, are these “individual solutions”? Are respondents simply too lazy to make their choice among the answer categories offered or do they have concrete ideas in mind? This was the research goal we pursued with our open-ended probing question. Table 2 presents by country the types of “individual solutions” that respondents think of regarding the division of labor between men and women.

The first two codes that we extracted from the open-ended answers offered by our respondents, *time availability* and *resource dependency*, refer to general rules which depend less on personal decisions and preferences of the family or the partners involved. *Time availability* connects the decision on household labor to the labor-force involvement. Respondents refer to the time resources of both partners. The division of household labor should take into consideration how much time is left after paid work (e.g. “It depends on the jobs the parents have, whether it is possible to work part-time”). This argumentation pattern is gender neutral. It also leaves – as a general rule – open, how the division of market labor is established.

*Resource dependency* is broader in that it also connects the decision about who might work outside of the home and who might stay at home to the earning

*Table 2* Answers to the category-selection probe for respondents who opted for the “individual solutions” category in the closed question in the different countries (in percent)

	Germany	Great Britain	United States	Mexico	Spain
Time availability	10	2	3	12	17
Resource dependency	21	16	8	11	13
Individual preferences	8	7	4	3	5
Individual abilities	0	8	2	5	3
Family/partners have to decide					
- no interference	3	10	11	1	1
- joint decision	3	7	11	13	8
- general	10	5	6	2	1
Situation dependency					
- financial necessities	14	9	13	15	17
- presence of one parent at home	8	9	8	6	6
Alternative possibilities	13	5	7	10	17
No generalization possible	26	38	29	34	33
Substantive rest category	6	7	15	12	6
Probe nonresponse	3	2	1	0	1
N	144	164	168	117	173

*Data source:* Web survey, Split 2; multiple coding possible for all categories except for no generalization possible, substantive rest category, and probe nonresponse; that is, figures do not add up to 100%.

potential of the partners. The person with the higher earning potential or career opportunity should work outside the home. As the citation “Well it could be the case where the mother could earn more income in her job than the father could and therefore it would be better for the mother to work than the father” reveals, this argumentation pattern is – in principle – again gender neutral. Admittedly, this argumentation pattern – and the same applies to *time availability* – can be used by traditional respondents, too, especially when they surmise that men will earn more than women anyhow in most cases. For time availability an additional caveat is necessary if the amount of labor-force participation of the woman is not reflecting her free will but has been kept low by the intervention of the man. As a consequence, it is not possible to unambiguously gauge the traditionality of respondents who opt for these categories.

*Time availability* is a frequent criterion in Spain (17%) and, to a smaller degree also in Mexico (12%) and Germany (10%), but it is rarely used in Great Britain and the United States (2% and 3%, respectively). *Resource dependency* as a criterion is clearly more popular than time availability in Germany (21%) and the two Anglo-Saxon countries (16% in Great Britain and 8% in the United States), and of nearly equal importance as time availability in Mexico and Spain (11% and 13%, respectively).

The code *individual preferences* captures when respondents refer to the partners' interests and preferences which should decide on the division of labor ("Because some women and men would rather stay home and take care of their house or their kids and some want to work").

Country differences with regard to *individual preferences* are not pronounced, ranging from 3% in Mexico to 8% in Germany. In general, both of these codes are of minor importance compared to *time availability* and *resource dependency*.

The code *individual abilities* reflects capabilities of the partners with regard to the job and household chores or childraising as the main decision criterion (e.g. "Every home situation is personal. It depends on which parent has the best career prospects and ability to support the family but also who would be the most suitable parent to take more responsibility raising the children"). *Individual abilities* have a similar importance as *individual preferences* in most countries, ranging from 0% to 8%. However, what is striking is the complete absence of this criterion in Germany.

A further important criterion for the decision on individual solutions is the idea that the *family/partners have to decide by themselves*. Respondents differ in their focus: *No interference* stresses that the society or other people in general have no right to intervene in this private decision (e.g. "Democracy allows individual freedom. The State has no place interfering in personal lives"). *Emphasis on joint family decisions* indicates that a consensus in the family should be reached which might involve engaging in compromises (e.g. "... if it is agreeable to both parents"). Respondents also made rather general statements, which are not pronounced enough to be classified into one of the two previous codes (e.g. "You cannot offer a solution for all. That has to be individually decided by the respective families").

Overall, in Great Britain and the United States, respondents are clearly more in favor of the family or the partners to decide on the division of household labor than in the other countries. Both countries belong to the liberal regime type where the state is not assumed to intervene in family life and does neither actively facilitate nor hinder the combination of family and work roles (by men and women). This kind of individualism is expressed with most vigor in the *no interference* category which holds any outside intervention (and maybe even advice) into family decisions to be illegitimate: 10% of the respondents in Great Britain and 11% in the United States share this stance compared to only 1-3% in the other three countries.

Respondents also took the situational context into account when responding to the probe. *Situation dependency – financial necessities* applies when the organization of the role division should be decided taking the financial necessities of a family into account, in particular whether a double income is needed to make ends meet (e.g. “Sometimes it is necessary for both parents to work in order to financially provide for their child and family. However, if it is possible to live comfortably with just one parent working, then it is up to the parents to decide how they want to raise their family”). *Financial necessities* come to mind quite frequently. Spanish respondents think of this aspect most often (17%). This does not come as a surprise, as Spain is one of the countries which were most severely hit by the financial crisis (beginning in 2007) and the web survey was conducted during its peak/aftermath in 2014. On the contrary, British respondents are the least frequent to mention this aspect (9%).

In contrast, we assigned the code *situation dependency – presence of one parent at home*, when respondents favor a model in which one person goes out to work and the other cares for the children and the household. Whether the man or the woman goes out to work or stays home is irrelevant – at least to most respondents (e.g. “I believe that pre-school children benefit most from having a parent care for them full-time, but it does not matter if it is father or mother”). The call for *the presence of one parent at home* is of moderate frequency and country differences are relatively small, ranging from 6% in Mexico and Spain to 9% in Great Britain.

Respondents also thought of *alternative possibilities* to fulfill the needs of the children that are not related to the allocation of work roles among the parents. Examples are the involvement of grandparents as well as privately or publicly organized daycare (e.g. “There are various support systems available within different families, so no particular hard rule can apply in all instances”).

Spanish and German respondents (17% and 13%, respectively) more often think of *alternative possibilities* (such as the involvement of grandparents or daycare) while in the Anglo-Saxon countries such a response is less frequent (5% in Great Britain and 7% in the United States). Contrary to our expectations, Mexicans are in-between.

*No generalization possible* was coded when respondents referred to individual differences in general without specifying any concrete criteria for the division of labor between both genders (e.g. “There is no right or wrong way, no one solution can suit every family”). Between one fourth (Germany) and more than one-third (Great Britain) of the respondents are coded into this category, thus referring to individual differences in general without specifying any concrete criteria. This could be an effect of web probing. Due to the web implementation, there is no possibility to spontaneously follow-up on answers that are not yet sufficiently clear.

The *substantive rest category* comprises answers that cannot be categorized into the substantive codes (e.g. “That again is freedom”) or are difficult to compre-

hend. Between 6% of the respondents in Germany and Spain and 15% in the United States give a response that we coded as *substantive rest category*. This means that we could not fit the response into our category scheme (and similar responses did not occur frequently enough to justify the addition of additional categories) or it was not sufficiently comprehensible to assign it unambiguously to one of the existing codes. This is unfortunately a weakness of web probing, namely, that it does not allow for a clarification of unclear statements made by respondents (a problem that could be easily solved by the interviewer in a cognitive interview; see Meitinger & Behr, 2016).

Finally, *probe nonresponse* includes explicit refusals, “don’t knows”, and answers such as “dddf”. This topic, however, is not affected by *probe nonresponse*; nearly all respondents try to give a substantive answer.

In addition to the general prevalence of response categories, we also conducted an analysis of gender differences with regard to these codes. However, there are hardly any consistent differences between men and women across all countries (see Table A8 in the Appendix). In most countries, however, women are more likely to refer to *situation dependency – financial necessities* and *alternative possibilities* than men. On the contrary, they are less likely to opt for the *no generalization possible* category than men.

## Discussion

In this paper, we demonstrated the usefulness of web probing when there is only one item to validate. We used the example of a new instrument in the ISSP module on “Family and Changing Gender Roles” (2012). The new instrument asks respondents to select one out of six role divisions between men and women when there are children at home. An “individual solutions” category was not added in the original ISSP questionnaire due to some concern that traditional respondents might use this category to avoid an overt disclosure of their traditional positions.

Our results show, however, that the “individual solutions” category is likely to be used by all kinds of respondents, not only the traditional ones. This was revealed both by the experimental quantitative and the qualitative data. The experiment showed that the addition of an “individual solutions” category to the response alternatives of the ISSP question was most attractive for less traditional respondents who would otherwise opt for the “both part-time” response alternative, and to those who would otherwise choose the “don’t know” category. The qualitative data from the web probing was likewise very informative, even though the single largest group of respondents in all countries referred to differences between individuals and families in a general way. Nevertheless, most respondents mentioned concrete criteria that should be used in families for coming to a decision on the optimal divi-

sion of labor between men and women. These criteria are mainly gender neutral, at least at face value. However, as mentioned above, as these criteria can also be used by traditional respondents (who take the inequality between both genders in the underlying conditions for granted), it is not possible to unambiguously infer the traditionality of respondents from these answers. Although the inclusion of an “individual solutions” category did not lead to the anticipated consequence (namely that it would attract mostly traditional respondents who did not want to explicitly express their position), it can nevertheless not be recommended for a regular survey that is not supplemented by web probing. This is because the selection of the “individual solutions” category cannot unambiguously be interpreted without the information from web probing and, thus, a clearly interpretable response would be missing for about half of the respondents.

Comparing the countries in our study, the majority of the criteria are of roughly the same importance in all countries or most of them. However, there are some noteworthy exceptions. In the two Anglo-Saxon countries (Great Britain and the United States), time availability was not mentioned frequently. In addition, in these two countries alternative possibilities of child care – outside the nuclear family – are less seen as a potential remedy to help decide on the role division between both partners. Instead, and in line with our hypothesis, in Great Britain and the United States, respondents make a point in that it is the family and the partners who have to decide this issue, and interference from outside of the family (in particular by the society at large) is seen as largely illegitimate.

In any case, the mixed-methods approach was crucial in assessing the consequences of adding an “individual solutions” category to the newly constructed ISSP item. We started with the quantitative ISSP data and compared it with the quantitative data of our web survey in order to establish whether it is possible to generalize results obtained from the latter to the former survey. Within the web survey, we then conducted a question experiment where the treatment group received an additional response category. Finally, this additional response category was probed and using the qualitative information obtained from the web survey the probe answers were coded and analyzed in a quantitative manner. The mixed-methods approach chosen allowed us to gain insights that we could not possibly have obtained by using a quantitative or qualitative method alone.

Several limitations of our study have to be mentioned. First, we used data based on non-probability online surveys. In order to tackle the issue whether we can use the web survey to shed light on the ISSP survey we compared the distribution of the central variable which was measured in the same way in one of the experimental splits and the ISSP survey. Nevertheless, we cannot exclude that results from probing could be somewhat different for the general population compared to the web survey.



Second, our experimental and probing data is limited to five countries and in these the highly developed countries are overrepresented compared to the ISSP survey. Only by replicating our study in additional countries in which the ISSP is conducted can we become more confident that our findings describe a general tendency in answer behavior and are not restricted to the countries we selected.

This paper does not inform on the more general question whether multiple-item measures are more adequate to measure gender-role attitudes than the single question we have analyzed. The evidence collected here is restricted to deciding in favor or against an inclusion of an “individual solution” category in the new ISSP question. In more general terms, while multi-item measures have clear advantages compared to a measure consisting of a single question (e.g. the possibility to employ data-analytic methods to establish equivalence across countries), there are also shortcomings with (existing) multi-items measures in large-scale comparative research. While most of the extant questions are concentrated on the role of the woman and might have a traditional slant, the construction of more balanced items which allow capturing egalitarian attitudes is also challenging, as there are a variety of possible egalitarian stances (Braun, 2008). In the end, the new ISSP measure was one attempt to bypass these shortcomings. At least in the area of gender-role attitudes, both question formats (multiple-item batteries and single questions) seem to have their merits (and weaknesses).

## References

- Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., ... Lavrakas, P. J. (2010). AAPOR report on online panels. *Public Opinion Quarterly*, *74*, 711–781.
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*, 287–311.
- Behr, D. (2015). Translating answers to open-ended survey questions in cross-cultural research: A case study on the interplay between translation, coding, and analysis. *Field Methods*, *27*, 248–299.
- Behr, D., & Braun, M. (2015). Satisfaction with the way democracy works: How respondents across countries understand the question. In P. B. Sztabinski, H. Domanski, & F. Sztabinski (Eds.), *Hopes and anxieties. Six waves of the European Social Survey* (pp. 121–138). Frankfurt am Main: Lang.
- Behr, D., Meitinger, K., Braun, M., & Kaczmirek, L. (2017). Web probing – Implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines), DOI: 10.15465/gesis-sg\_en\_023.
- Benítez, I., & Padilla, J.-L. (2014). Analysis of nonequivalent assessments across different linguistic groups using a mixed methods approach: Understanding the causes of differential item functioning by cognitive interviewing. *Journal of Mixed Methods Research*, *8*, 52–68.

- Benítez, I., Padilla, J.-L., van de Vijver, F., & Cuevas, A. (2018). What cognitive interviews tell us about bias in cross-cultural research: An illustration using quality-of-life items. *Field Methods, 30*, 277-294.
- Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing the housework? Trends in the gender division of household labor. *Social Forces, 79*, 191-228.
- Bittman, M., England, P., Sayer, L. C., Folbre, N., & Natheson, G. (2003). When does gender trump money? Bargaining the time in household work. *American Journal of Sociology, 109*, 186-214.
- Blasius, J., & Thiessen, V. (2006). Assessing data quality and construct comparability in cross-national surveys. *European Sociological Review, 22*, 229-242.
- Braun, M. (2008). Using egalitarian items to measure men's and women's family roles. *Sex Roles, 59*, 644-656.
- Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *American Journal of Sociology, 100*, 652-688.
- Brogan, D., & Kutner, N.G. (1976). Measuring sex-role orientation: a normative approach. *Journal of Marriage and the Family, 38*, 31-40.
- Clogg, C. C. (1984). Some statistical models for analyzing why surveys disagree. In Turner, C. F., & Martin, E. (Eds.), *Surveying subjective phenomena. Volume 2* (pp. 319-366). New York: Russell Sage.
- Creswell, J. W. (2014). *Research design. Qualitative, quantitative, and mixed methods approaches*. 4<sup>th</sup> edition. Los Angeles: Sage.
- Davidov, E., Dülmer, H., Schlüter, E., Schmidt, P. & Meuleman, B. (2012). Using a multilevel structural equation modeling approach to explain cross-cultural measurement noninvariance. *Journal of Cross-Cultural Psychology, 43*, 558-575.
- Edlund, J., & Öun, I. (2016). Who should work and who should care? Attitudes towards the desirable division of labour between mothers and fathers in five European countries. *Acta Sociologica, 59*, 151-169.
- Fitzgerald, R., Widdop, S., Gray, M., & Collins, D. (2011). Identifying sources of error in cross-national questionnaires: Application of an error source typology to cognitive interview data. *Journal of Official Statistics, 27*, 569-599.
- Glick, P. & Fiske, S. T. (1997). Hostile and benevolent sexism. Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly, 21*, 119-135.
- Harkness, J. A. (2003). Questionnaire translation. In Harkness, J. A., van de Vijver, F. J., & Mohler P. (Eds.), *Cross-cultural survey methods* (pp. 35-56). New York: Wiley.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409-426.
- ISSP Research Group (2016). International Social Survey Programme: Family and Changing Gender Roles IV - ISSP 2012. GESIS Data Archive, Cologne. ZA5900 Data file Version 4.0.0, DOI: 10.4232/1.12661.
- Kalleberg, A. L. & Rosenfeld, R. A. (1990). Work in the family and in the labor market: A cross-national, reciprocal analysis. *Journal of Marriage and the Family, 52*, 331-346.
- King, L. A., & King, D. W. (1997). Sex-role egalitarianism scale. Development, psychometric properties, and recommendations for future research. *Psychology of Women Quarterly, 21*, 71-87.
- Knight, C. R., & Brinton, M. (2017). One egalitarianism or several? Two decades of gender-role attitude change in Europe. *American Journal of Sociology, 122*, 1485-1532.

- Luyt, R. (2012). A framework for mixing methods in quantitative measurement development, validation, and revision: a case study. *Journal of Mixed Methods Research*, 6, 294–316.
- Meitinger, K. (2017). Necessary but insufficient: Why measurement invariance tests need online probing as a complementary tool. *Public Opinion Quarterly*, 81, 447–472.
- Meitinger, K. & Behr, D. (2016). Comparing cognitive interviewing and online probing: Do they find similar results? *Field Methods*, 28, 363–380.
- Meuleman, B. & Billiet, J. (2009). A Monte Carlo sample size study: How many countries are needed for accurate multilevel SEM? *Survey Research Methods*, 3, 45–58.
- Miller, K., Fitzgerald, R., Padilla, J.-L., Willson, S., Widdop, S., Caspar, R., ... Schoua-Glusberg, A. (2011). Design and analysis of cognitive interviews for comparative multinational testing. *Field Methods*, 23, 379–396.
- Scholz, E., Jutz, R., Edlund, J., Öun, I., & Braun, M. (2014). ISSP 2012. Family and Changing Gender Roles IV. *GESIS-Technical Reports 2014/19*. Köln: GESIS. <https://www.gesis.org/issp/modules/issp-modules-by-topic/family-and-changing-gender-roles/2012/>
- Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, 68, 199–214.
- Thrasher, J. F., Quah, A. C., Dominick, G., Borland, R., Driezen, P., Awang, R., ... Boado, M. (2011). Using cognitive interviewing and behavioral coding to determine measurement equivalence across linguistic and cultural groups: An example from the International Tobacco Control Policy Evaluation Project. *Field Methods*, 23, 439–460.
- van de Vijver, F. J. R., & Chasiotis, A. (2010). Making methods meet: Mixed designs in cross-cultural research. In Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P., ... Smith, T. (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 455–473). Hoboken, NJ: Wiley.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks: Sage.
- Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly*, 79, 359–395.

## Appendix

*Table A1* Quota assignment in the web survey

Age	Gender	Education	Germany (539)	Great Britain (534)	United States (540)	Spain (531)	Mexico (545)
18-30	Male	High	8.53	8.24	8.15	8.29	8.99
18-30	Male	Low	8.16	8.24	8.15	8.29	8.26
18-30	Female	High	8.91	8.61	8.15	8.29	8.62
18-30	Female	Low	8.35	8.80	8.52	8.29	8.26
31-50	Male	High	8.16	8.24	8.33	8.29	8.62
31-50	Male	Low	8.35	8.24	8.52	8.66	8.07
31-50	Female	High	8.16	8.24	8.15	8.29	8.26
31-50	Female	Low	8.35	8.24	8.52	8.29	8.07
51-65	Male	High	8.16	8.24	8.33	8.47	8.07
51-65	Male	Low	8.16	8.24	8.33	8.29	8.07
51-65	Female	High	8.53	8.33	8.70	8.29	8.44
51-65	Female	Low	8.16	8.24	8.15	8.29	8.26
			100%	100%	100%	100%	100%

*Table A2* Preferred division of labor for ISSP question in the different countries (in percent)

	Germany		Great Britain		United States		Mexico		Spain	
	ISSP	Web	ISSP	Web	ISSP	Web	ISSP	Web	ISSP	Web
Mother at home, father full-time	20	20	34	27	29	27	49	22	24	4
Mother part-time, father full-time	44	35	38	28	32	23	23	50	39	21
Both full-time	10	10	4	10	9	23	7	11	11	22
Both part-time	13	23	4	13	5	6	16	15	18	45
Father part-time or at home, mother full-time	1	1	0	1	1	1	3	0	1	0
Don't know	13	13	20	20	25	20	3	2	7	8
N	1,766	275	950	281	1,302	266	1,527	253	2,595	268

*Data source:* ISSP 2012; Web survey, split 1; original categories “father part-time and mother full-time” and “father at home and mother full-time” collapsed.

*Table A3* Preferred division of labor for ISSP question in the different countries, separately for male and female respondents (in percent)

	Germany		Great Britain		United States		Mexico		Spain											
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female										
	ISSP	Web	ISSP	Web	ISSP	Web	ISSP	Web	ISSP	Web										
Mother at home, father full time	24	17	14	38	27	30	27	34	29	25	26	48	26	50	17	29	4	20	4	
Mother part time, father full time	41	29	47	41	35	30	41	26	30	20	33	26	23	53	22	46	37	18	41	24
Both full time	10	12	9	7	4	11	4	10	8	26	9	20	7	9	7	15	12	25	11	19
Both part time	11	21	14	24	4	11	5	15	3	6	6	6	16	11	15	20	15	43	20	48
Father part time or at home, mother full time	0	1	1	0	0	3	0	0	1	1	0	1	2	0	3	0	1	0	1	0
Don't know	14	12	12	15	20	18	20	22	24	18	26	21	4	1	3	2	7	11	7	6
<i>N</i>	857	145	909	130	438	152	512	129	594	129	708	137	727	138	796	115	1,271	134	1,378	134

*Data source:* ISSP 2012; Web survey, split 1.

*Table A4* Preferred division of labor dependent on the presence of an “individual-solutions” category in the different countries, separately for male and female respondents (in percent)

	Germany		Great Britain		United States		Mexico		Spain												
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female											
	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2											
Mother at home, father full time	26	13	14	10	27	14	27	13	29	14	26	9	26	19	17	14	4	4	4	4	2
Mother part time, father full time	29	24	41	21	30	13	26	9	20	12	26	10	53	30	46	28	18	8	24	12	12
Both full time	12	6	7	3	11	7	10	3	26	14	20	6	9	10	15	7	25	6	19	7	7
Both part time	21	7	24	6	11	3	15	4	6	4	6	1	11	4	20	5	43	17	48	12	12
Father part time or at home, mother full time	1	0	0	0	3	1	0	0	1	1	1	0	0	1	0	1	0	1	0	0	0
Individual solutions	-	49	-	59	-	61	-	68	-	50	-	73	-	36	-	44	-	65	-	67	-
Don't know	12	2	15	1	18	2	22	3	18	5	21	1	1	1	2	1	11	0	6	0	0
<i>N</i>	145	122	130	142	152	112	129	141	129	140	137	134	138	135	115	157	134	133	134	130	130

*Data source:* Web survey; split 1: original ISSP version, split 2: “individual-solutions” category added.

*Table A5* Preferred division of labor dependent on the presence of an “individual-solutions” category in the different countries (in percent; calculation excluding “don’t know” and “individual solutions” categories)

	Germany		Great Britain		United States		Mexico		Spain	
	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2	Split 1	Split 2
Mother at home, father full-time	23	26	34	42	34	33	22	27	4	9
Mother part-time, father full-time	40	50	35	31	29	32	51	49	23	28
Both full-time	11	9	13	15	29	28	12	14	24	19
Both part-time	26	15	16	11	8	6	15	8	49	43
Father part-time or at home, mother full-time	0	0	2	1	1	2	0	2	0	1
<i>N</i>	239	117	224	83	214	98	249	171	246	90

*Data source:* Web survey; split 1: original ISSP version, split 2: “individual-solutions” category added; original categories “father part-time and mother full-time” and “father at home and mother full-time” collapsed.

*Table A6* Preferred division of labor dependent on the presence of an “individual-solutions” category in the different countries, separately for male and female respondents (in percent; calculation excluding “don’t know” and “individual solutions”)

	Germany		Great Britain		United States		Mexico		Spain											
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female										
	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2	Split 1 2										
Mother at home, father full time	29	27	16	25	33	38	35	46	36	32	34	29	17	26	5	11	4	7		
Mother part time, father full time	33	48	48	53	37	33	33	29	25	27	33	40	47	47	51	20	21	25	35	
Both full time	13	12	8	7	13	19	13	10	31	30	26	23	9	15	13	28	17	20	21	
Both part time	24	13	28	16	14	7	19	15	8	8	7	3	11	6	20	9	48	49	51	37
Father part time or at home, mother full time	1	0	0	0	3	2	0	0	1	3	1	0	2	0	1	0	2	0	0	0
<i>N</i>	128	60	111	57	124	42	100	41	106	63	108	35	136	85	113	86	120	47	126	43

*Data source:* Web survey; split 1: original ISSP version, split 2: “individual-solutions” category added.



*Table A7* Choice of individual-solution category in different social groups (in percent)

	Germany	Great Britain	United States	Mexico	Spain
Women	59	68	73	44	67
Men	49	61	50	36	65
Average year of birth (individual-solutions category not selected)	1974	1974	1973	1975	1976
Average year of birth (individual-solutions category selected)	1971	1970	1973	1974	1973
Married	52	61	60	35	67
Not married	56	67	63	46	66
Full-time employed	54	63	57	40	66
Part-time employed	56	64	51	42	70
Not in employment	55	68	69	40	64
Partner full-time employed	52	63	63	40	71
Partner part-time employed	47	68	50	43	48
Partner not in employment	56	64	65	21	65
Children yes	54	64	62	40	66
Children no	55	66	61	40	65

*Data source:* Web survey, Split 2; multiple coding possible for all categories but no generalization possible, other answers, and probe nonresponse, i.e. figures do not add up to 100%.

*Table A8* Answers to category-selection probe for respondents who opted for “individual-solution category” for closed question in the different countries, separately for male and female respondents (in percent)

	Germany		Great Britain		United States		Mexico		Spain	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Time availability	7	12	1	3	1	4	13	12	15	18
Resource dependency	17	24	16	17	9	7	8	13	13	14
Individual preferences	2	13	2	11	3	4	2	4	7	3
Individual capabilities	0	0	10	6	3	2	4	6	6	1
Family/partners have to decide										
- general	5	13	3	6	6	6	2	1	0	2
- no interference	3	2	9	11	6	15	2	0	0	1
- joint decision	3	2	7	6	9	12	19	9	8	8
Situation dependency										
- financial necessities	10	17	9	9	9	15	10	19	14	21
- presence of one parent at home	8	8	9	9	9	7	4	7	8	5
Alternative possibilities	12	14	1	8	7	7	6	13	15	20
No generalization possible	32	23	44	33	33	26	35	33	33	33
Other answers	13	0	6	7	17	14	13	12	7	5
Probe nonresponse	2	5	3	2	1	0	0	0	0	1
<i>N</i>	60	84	68	96	70	98	48	69	86	87

*Data source:* Web survey, Split 2; multiple coding possible for all categories but no generalization possible, other answers, and probe nonresponse, i.e. figures do not add up to 100%