HHS Public Access

Author manuscript

*J Fam Stud*. Author manuscript; available in PMC 2023 March 01.

Published in final edited form as:

*J Fam Stud*. 2022 ; 28(4): 1355–1376. doi:10.1080/13229400.2020.1825229.

# Measuring Maternal Gatekeeping: A Rasch Analysis of the Parental Regulation Inventory

**Lauren E. Altenburger**[1], **P. Cristian Gugiu**[2], **Sarah J. Schoppe-Sullivan**[3]

[1]Department of Human Development and Family Studies, Pennsylvania State University – Shenango, Sharon, PA 16146

[2]Independent Researcher

[3]Department of Psychology, The Ohio State University, Columbus, OH 43210

## Abstract

To understand factors that may influence father involvement, researchers have increasingly considered maternal gatekeeping, or the extent to which mothers might attempt to regulate (i.e., encourage, discourage) fathers' involvement in childrearing. Although several theoretical models of maternal gatekeeping have been advanced in recent years, maternal gatekeeping measurement has lagged significantly behind developments in gatekeeping theory. Rasch analysis offers a useful framework for conducting item-level analyses to evaluate measurement validity and identify areas of improvement for measurement scales. In the present study, Rasch analysis techniques were implemented to 1) illustrate how modern psychometric methods can be applied to validate measures in family psychology and 2) examine the validity of the Parental Regulation Inventory, a commonly used maternal gatekeeping measure (PRI; Van Egeren, 2000). Results indicated that the PRI exhibited adequate construct validity; however, measurement could be improved by including additional items on the PRI subscales. In particular, Rasch analyses indicated floor effects on fathers' reports of maternal gate closing, floor and ceiling effects on fathers' reports of maternal gate opening, and floor and ceiling effects on fathers' reports of maternal communication at 3- and 9- months postpartum. Recommendations for improving maternal gatekeeping measurement and implications for maternal gatekeeping theory are discussed.

### Keywords

maternal gatekeeping; father involvement; Parental Regulation Inventory; Rasch analysis; measurement

Theoretical models describing influences on fathers' parenting have pointed to the critical roles of mothers (Cabrera et al., 2014; Cummings et al., 2004; Doherty et al., 1998). Indeed, as romantic partners and coparents, mothers may influence the quantity and quality of fathers' parenting (Fagan & Palkovitz, 2011). In an effort to better understand how fathers' relationships with children's mothers may affect their involvement in parenting, researchers

lqa5222@psu.edu .

The authors declare they have no conflict of interest.

have considered maternal gatekeeping. As originally conceptualized, maternal gatekeeping consists of attitudes and behaviors that serve to limit fathers' involvement with children (Allen & Hawkins, 1999). More recently, scholars have expanded upon the original Allen and Hawkins (1999) gatekeeping framework to account for the possibility that mothers might facilitate, encourage, discourage, or control fathers' involvement in childrearing (Cannon et al., 2008; Fagan & Cherson, 2017; Puhlman & Pasley, 2013). Although multiple conceptual models of maternal gatekeeping have emerged in recent years, measurement of maternal gatekeeping has lagged behind theory development.

Several survey measures of maternal gatekeeping have been used in a handful of empirical studies (i.e., Fagan & Barnett, 2003; Meteyer & Perry-Jenkins, 2010; Puhlman & Pasley, 2017; Schoppe-Sullivan et al., 2008). However, there is no consensus among researchers on which measure best characterizes maternal gatekeeping. The earliest measures of maternal gatekeeping focused largely on assessing mothers' *beliefs* about the importance of fathers and the division of household and child care responsibilities (Allen & Hawkins, 1999; Fagan & Barnett, 2003). However, critics argued that measuring mothers' gatekeeping attitudes or division of labor preferences alone does not capture a gatekeeping process. To shift the focus more strongly on mothers' gatekeeping *behaviours* – encouraging or discouraging actions directed toward the father – other measures of maternal gatekeeping were used. One unpublished survey measure that has been included in several studies of families with young children is the Parental Regulation Inventory (PRI; Van Egeren, 2000). Although originally developed to measure regulation of a coparenting partner's involvement in childrearing, researchers have used a subset of items taken from the PRI (Van Egeren, 2000) to measure maternal *gate opening* (i.e., tell your baby's father how happy he makes your baby) and *gate closing* behaviours (i.e., take over and do it your own way) (Schoppe-Sullivan et al., 2008). The PRI provides a promising step in advancing maternal gatekeeping theory and research, as it offers an advantage over other measures of maternal gatekeeping in its focus on parents' perceptions of maternal gatekeeping *behaviours*—which are more central to the gatekeeping construct— rather than maternal gatekeeping *attitudes* (Schoppe-Sullivan et al., 2015).

Scholars using a subset of items from the PRI (Van Egeren, 2000) have reported associations between perceptions of maternal gatekeeping behaviour and father involvement (Altenburger et al., 2017; Schoppe-Sullivan et al., 2008; Zvara et al., 2013). Although these analyses provide important insight into the implications of maternal gatekeeping for father involvement, a thorough item-level examination of the PRI as a measure of maternal gatekeeping has yet to be conducted. Rasch analysis provides a useful framework for assessing both the quality of survey measures, as well as identifying directions for improving them. Although the benefits of Rasch analysis are numerous (Bond & Fox, 2007; Boone et al., 2014), and several fields, including medicine (Duncan et al., 2003), education (Boone et al., 2011), and clinical psychology (Elliot et al., 2006) have capitalized on its features to improve the measurement of targeted constructs, it has rarely been applied in family psychology research to improve the quality of survey instruments and advance theory. Thus, a primary goal of this study was to illustrate the utility of Rasch analysis for survey development in the field of family psychology. A secondary goal of this study was to evaluate the validity of the PRI using Rasch analysis techniques (Van Egeren, 2000). Rasch analysis is a powerful analytic framework that has many practical implications for

evaluating and improving maternal gatekeeping measurement. The validity of the PRI will be determined by considering the internal validity of maternal gatekeeping items, examining whether items capture the full range of maternal gatekeeping levels in a sample of new parents (i.e., very low gate closing to very high gate closing), and evaluating the extent to which floor and ceiling effects were exhibited. Implications for maternal gatekeeping theory and improving maternal gatekeeping measurement will be discussed.

## Rasch Analysis for Instrument Validation

Strong measurement in family research is important for the advancement of theoretical models. A key component of quantitative research is defining constructs of interest and operationalizing them on a measurement tool such as a survey instrument. Typically, in illustrating the reliability and validity of survey instruments, researchers must justify that a particular measure is an acceptable assessment of the underlying construct of interest. It is common for researchers to do this by referring to previously published researh that has used the scale, the original scale developers' work, or reporting internal consistency estimates, such as Cronbach's alphas. With the advancement of psychometric theory and research, however, there are many analytic tools readily available to researchers that can be implemented to better evaluate and improve the quality of survey instruments. Rasch analysis offers a number of benefits for family researchers interested in evaluating surveys, including the following: 1) item difficulty and person measures can be applied to assess the validity of survey instruments, 2) Wright maps can be used to identify directions for survey improvement, and 3) bias in individual survey items can be detected.

### Assessing Validity of Survey Instrument

Rasch analysis provides a framework that researchers can leverage to assess the validity of survey measures. Underlying Rasch analysis is the assumption that participants vary in the degree to which they exhibit a particular behavior or embody a trait. That is, items are thought to fall at different intervals along a ruler (Gordon, 2015). Applied to maternal gatekeeping, researchers would expect the presence of maternal gatekeeping in families to exist on a continuum of low frequency occurrences to high frequency occurrences. For example, some participants might consistently report very low levels of maternal gatekeeping (i.e., gatekeeping is virtually absent in their family), whereas other participants might report very high levels of maternal gatekeeping (i.e., gatekeeping occurs several times a day). High-quality survey instruments measure a construct by including items that appropriately target participants' position on a continuum of interest and differentiate between participants who embody varying levels of the construct (Brown, 2014). Rasch analysis produces item difficulty estimates and person measures which enable researchers to assess how well items target the participants.

**Item Difficulty—**Psychometricians using Rasch analysis expect a well-designed survey instrument to include items that vary in "item difficulty," or the likelihood that participants will agree with an item (Boone et al., 2014). On a survey, an extremely difficult item means few participants agree with the item. Applied to maternal gatekeeping, the following item might have high item difficulty: "On a scale of 1 to 6 (1 = *never*, 6 = *several times per*

*day*), how often does the following occur, 'My child's mother undermines my parenting in public.'" This is an overt example of maternal gatekeeping, in which the mother is directly undermining the father in a public setting. If, in fact, few fathers select higher frequencies (i.e., 5 and 6) and several fathers select lower frequencies (i.e., 1 and 2), then this item would have a high item difficulty level. In contrast, an item that has a lower difficulty level might be: "On a scale of 1 to 6 (1= *never*; 6 = *several times per day*), how often does the following occur, 'My child's mother shows nonverbal dislike (i.e., rolls eyes) when I do something with the child that she does not like.'" If several fathers select higher frequencies (i.e., 5 and 6) and few fathers select lower frequencies (i.e., 1 and 2), then this item would have a low item difficulty level. The item is easy for fathers to endorse. In designing a high-quality survey, researchers strive to develop items that range in item difficulty level—some items are expected to be easier for participants to endorse and other items are expected to be harder to endorse. Because subtle behaviors like eye-rolling are likely to occur more frequently in families, researchers may hypothesize that this item would have a lower difficulty level than an item that asks fathers about how often their baby's mother undermines them in public. Once the survey is administered, researchers can apply Rasch analysis techniques to evaluate whether the item difficulty levels align with the researcher's expectations (Boone et al., 2014).

**Person Measures—**In addition to item difficulty, Rasch analysis allows researchers to calculate "person measures," which indicate the extent to which participants embody a latent trait of interest (i.e., maternal gatekeeping). Rasch analysis produces a single person measure for each participant. A high person measure means that a participant very frequently endorsed survey items, including survey items that have a high item difficulty level. In the context of maternal gatekeeping, a father with a high person measure would have a greater likelihood of endorsing an item with high item difficulty than a father with a low person measure. For example, a father with a high person measure would be more likely to select a high value (i.e., 5 or 6) on the following item: "On a scale of 1 to 6 (1 = never, 6 = several times per day), how often does the following occur, 'My child's mother undermines my parenting in public.'" In contrast, a father with a low person measure may select a lower value on this item (i.e., 1 or 2). According to Rasch analysis, to measure a person well, the researcher must have items that have a lower item difficulty than the participant's person measure, as well as items that have a higher item difficulty than the participant's person measure. It is only by "sandwiching" the person between items that vary in difficulty that person measures can be reliably estimated (Bond & Fox, 2007). Each participant is "sandwiched" well if there are items that are easy for the participant to endorse (i.e., low item difficulty) and items that are hard for the participant to endorse (i.e., high item difficulty).

Item difficulty and person measures can be used together to evaluate a survey. A low-quality survey would include only items that have low item difficulty (i.e., everyone endorses) or only items that have high item difficulty (i.e., no one endorses). For example, if all participants indicate that a behaviour occurs "never" on survey items, there is no variability among participants. Perhaps the construct of interest truly never occurs in the sample. Alternatively, the behaviour may not be reported because the survey items have an item

difficulty that is too high. Items with a lower difficulty level should be developed. If the survey is administered again and participants are able to agree with items that have a lower difficulty level, then the researcher may conclude that the behaviour of interest occurs at lower than expected levels in their sample. Rasch analysis can aid family psychology researchers in identifying which items are redundant or not targeting participants well. No other analytic tool can provide this depth of information about how well items are functioning and how well participants are being measured.

## Using Wright Maps for Survey Development

If the data fit the Rasch model well, transformed scores follow a monotonically increasing pattern of item difficulty and person measures, which are on a common scale of "logit" units (Linacre & Wright, 1989). All Rasch analysis results are presented in logits, which is short for log odds units and allows researchers to make meaningful conclusions about the quality of their survey instrument (Boone et al., 2014). Because person measures and item difficulty are both on the same equal-interval logit scale, persons can be compared to other persons (i.e., person X reported a higher level of maternal gatekeeping than person Y) and items can be compared to other items (i.e., item 31 was easier for participants to agree with than item 10). Additionally, items and persons can be compared to each other (i.e., person X has a higher probability of agreeing with item 10 than person Y).

The Wright Map is a person-item map that can be produced in WINSTEPS software. Items are plotted on a scale in order from the item that has the lowest difficulty level (i.e., easy for participants to endorse) to the item that has the highest difficulty level (i.e., challenging for participants to endorse). People are also plotted on the same scale in order from persons with low person measure levels (i.e., exhibit low levels of latent trait of interest) to high person measure levels (i.e., exhibit high levels of latent trait of interest). Researchers should apply theory to predict what the item ordering might be and evaluate if the item ordering on the Wright map is consistent with theory. However, in the presence of little conceptual or empirical work, the item ordering could be used to refine a construct's operationalization (Gordon, 2015). Within a Rasch framework, content and construct validity can be established by assessing whether the items address the intended latent variable (content validity) and whether the item difficulty hierarchy aligns with theory (construct validity) (Baghaei, 2008).

A person is measured well when their position on the latent continuum is in between items that are easier for that person to agree with and items that are more challenging for participants to endorse (i.e., sandwiched). Items, in turn, are considered high in quality when there are several people with person measure levels near item difficulty levels. In some cases, large error estimates contribute to gaps in the instrument or sample, which might indicate that items are low in quality or people are not adequately measured. In order to improve a measure, it is beneficial for researchers to carefully consider which people and items are not well-targeted along the continuum (Bond & Fox, 2007). Because Wright maps provide the researcher with a visual diagram that places person measures and item difficulty levels on the same scale, the researcher can easily identify gaps along the continuum in which people are not measured well.

### Detecting Individual Item Bias

A strong measure comprises individual items that contribute in a meaningful way to the latent construct (Bond & Fox, 2007). Rasch statistics enable researchers to determine the degree to which items on an instrument measure the latent construct of interest. For example, Rasch fit statistics provide information about how well individual items or person responses cohere to the expectations of the Rasch model. When fit is acceptable, easy items are endorsed by a greater number of persons than difficult items. Additionally, persons with higher levels of the construct of interest (i.e., maternal gatekeeping) are more likely to endorse items that are difficult to agree with compared to persons with lower levels of the construct of interest.

## Rasch Analysis for Family Psychology Research

The Rasch model has taken root as a tool to evaluate the quality of various survey measures. In family psychology, some studies have applied Rasch analysis to calculate person measures for use in parametric analyses (Coley et al., 2011; Elliot et al., 2016; Phillipson & McFarland, 2016). Emerging research has also applied Rasch techniques to develop and validate a measure of parental involvement in children's elementary studies (Gugiu et al., 2019) and evaluate a measure of romantic relationship quality (Fowers et al., 2016). However, no study has used Rasch analysis to validate a measure of maternal gatekeeping. Rasch analysis offers family psychology researchers many resources for examining the precision and validity of a survey. In adopting a Rasch approach to survey development, family psychology researchers are challenged to think about items as falling at different intervals along a continuum. This requires researchers to connect items to the construct's underlying theory and think of items that might vary in difficulty. The feedback provided by Rasch analyses can be used to refine the underlying theory and how constructs of interest are operationalized.

## The Present Study

The primary aim of this study is twofold: 1) illustrate the utility of Rasch analysis for survey development in the field of family psychology and 2) examine the validity of the PRI by conducting a Rasch analysis to examine the quality of individual items, evaluate the extent to which floor and ceiling effects were exhibited, and identify directions for improvement. The present analysis focused exclusively on fathers' reports of maternal gatekeeping, as fathers' perceptions of maternal gatekeeping are arguably most important for fathers' subsequent levels of involvement (Altenburger et al., 2018). This study represents a preliminary attempt to increase knowledge about the strengths and weaknesses of the PRI, as well as to improve maternal gatekeeping measurement and theory. Moreover, this study provides an example to family researchers by illustrating the utility of Rasch analysis for improving family measurement and, in turn, theoretical conceptualizations. Although prior research has used various items from the PRI to measure maternal gatekeeping (Schoppe-Sullivan et al., 2015, 2008; Zvara et al., 2013), no prior study has conducted a Rasch analysis to evaluate its validity. Without a detailed item-level evaluation of maternal gatekeeping measurement, advancements in maternal gatekeeping theory may be hindered.

# Method

## Participants

Study participants were drawn from a longitudinal study of 182 fathers in different-sex, dual-earner relationships in which both parents were transitioning to parenthood in 2008 – 2009 and residing in a large, Midwestern U.S. city and surrounding area. Recruitment occurred at childbirth education classes, pregnancy health centers, and through the use of advertisements posted online, at doctors' offices, and in newspapers. Eligible participants were required to be married or cohabiting, 18 years of age or older, expecting their first biological child, able to read and speak English, currently employed full-time and planning to return to work postpartum, and planning to stay in the geographic area for at least one year. To comply with procedures approved by the University's Behavioral and Social Sciences Institutional Review Board, informed consent was obtained from each partner at each phase of the study.

The median level of education was a bachelor's degree and 65% of fathers reported having at least this education level. Median annual household income was $79,500. The average age of fathers was 30.20 years (*SD* = 4.81; *Range* = 18 – 50), and 86% identified as White, 7% as Black, 3% as Asian, 2% Hispanic, 4% as other races, and 1% as multi-racial. Eighty-five percent of fathers were married.

## Measures

**Maternal gatekeeping.—**Fathers reported maternal gatekeeping at 3- and 9- months postpartum using 35 items from the Parental Regulation Inventory (PRI; Van Egeren, 2000), which asked fathers to report the frequency (1 = *never*, 6 = *several times per day*) with which mothers encourage (i.e., "How often does your baby's mother do the following to encourage you to be involved in child care and with your baby, including feeding, play, and emotional support?") or disapprove of fathers' involvement in childrearing (i.e., "When you do something that your baby's mother doesn't approve of regarding child care or with your baby, how often does she do the following?").

## Analytic Technique

**Missing data analysis.—**Prior to exploratory factor analyses, missing data analyses were conducted on maternal gatekeeping variables. Examination of the usable missing values of interest at 3 months postpartum revealed a modest percentage of missingness ranging from 5.5% to 7.1%. Examination of the usable missing values at 9 months postpartum revealed a moderate percentage of missingness ranging from 17.0% to 18.7%. Missing values were imputed for maternal gatekeeping items using multiple imputation in IBM SPSS Statistical Package Version 25 to reduce bias associated with listwise deletion (see Tabachnick & Fidell, 2007). Multiple imputation was conducted separately for data at 3- and 9-months postpartum.

Given the relatively modest percentage of missing data, 5 data sets were imputed at 3- and 9-months postpartum, with higher than 96% efficiency. Little's Missing Completely at Random (MCAR) test was nonsignificant at 3-months ($\chi^2(1171) = 1205.5$, $p = .24$) and

9-months postpartum ($\chi^2(2116) = 2197.1$, $p = .11$). When Little's MCAR is nonsignificant, multiple imputations can estimate within 1% of the true value even when up to 50% of values included in the estimation model are missing (Scheffer, 2002). Results of multiply imputed data sets are often combined to provide a final estimate that has incorporated these data sets (Gugiu et al., 2010). However, statistical methods for aggregating values across imputed data sets have not been developed for exploratory factor analysis in commonly used statistical packages. Prior research has averaged imputed values (Jensen & Shafer, 2013) or used the median value across imputed data sets (Altenburger et al., 2017) to create a combined, imputed data set and conduct analyses. The "maximum median difference" and the "maximum variance difference" among imputed data sets are provided to confirm imputed values did not vary substantially between imputed data sets at 3- and 9-months postpartum. In the present sample, the differences between the median and variance of the five imputed data sets were minimal. The average median difference across imputed data sets was 0.03 at 3- months and .10 at 9-months postpartum. The average variance difference across imputed data set was 0.06 at 3-months and 0.14 at 9-months postpartum. Thus, the median values across all imputed data were used to obtain a single value for each variable in an aggregate dataset at 3- and 9-months postpartum.

**Exploratory factor analyses.—**Multiple theoretical conceptualizations of maternal gatekeeping exist (Allen & Hawkins, 1999; Fagan & Cherson, 2017; Puhlman & Pasley, 2013; Schoppe-Sullivan et al., 2008), yet no prior study has conducted an exploratory factor analysis on the Parental Regulation Inventory. Exploratory factor analyses were conducted in SAS version 9.4 on the Spearman 's rank correlation matrix of items to establish the dimensionality of maternal gatekeeping prior to Rasch analyses. Factor structures were determined through an iterative process of examining parallel and exploratory factor analysis results at 3- and 9-months postpartum (Glorfeld, 1995; Pett et al., 2003). Sample size requirements for EFA depend on a variety of data properties, including the size of the factor loadings and the extent to which factors are overdetermined (i.e., 5 variables on each factor). Monte Carlo literature has examined sample size conditions necessary for detecting stable factor structures across various data properties (Guadagnoli & Velicer, 1988). Results indicated that when the average factor loadings are near .60 and there are 3 factors with a sufficient number of factor loadings, the Kappa value (or measure of agreement) between the sample component patterns and population patterns is 1.0 for a sample size between 150 and 200. When the average factor loadings are near .40, the Kappa value is between .80 and .87 for a sample size between 150 and 200, respectively. Kappa values greater than .75 represent excellent agreement beyond chance.

**Rasch analysis.—**Rasch modeling was used to assess the quality of the PRI using WINSTEPS 4.5.0 software (JM Linacre, Beaverton, OR). To ensure Rasch analysis parameter estimates are unbiased, data should be fit to the Rasch model (Boone et al., 2014). The INFIT and OUTFIT MNSQ statistics were evaluated to determine item fit. Items that had a value below .5 or greater than 1.5 were flagged for further review (Linacre, 2020). To confidently interpret Rasch analysis estimates, misfitting items could be deleted. Alternatively, consistent with recommendations in the field (Linacre, 2002, 2020), model fit can be improved by iteratively evaluating outlier responses. Following these steps preserves

items and reduces bias in person measure estimates rather than introducing error (Gugiu et al., 2019; Wright, 1997; Wright & Stone, 1979). In our analysis, outlier responses with z-scores greater than |3.00| were set to missing on flagged items. To ensure item measures were not significantly influenced by setting misfitting responses to missing, sensitivity analyses were examined. Item measures were cross-plotted with the inclusion and exclusion of missing responses on the flagged item(s). If item measures fall along a straight line (within a 95% confidence interval), then the strategy for handling misfitting items does not bias item measures (Boone et al., 2014; Linacre, 2020; Persch et al., 2015).

Model fit was further established by examining whether there were negative item point-measure correlations or disordered response categories. Rasch reliability estimates were also computed to assess person and item measure reliability. Both reliability statistics reflect the extent to which rank-order and spacing between person and item measures would remain stable in study replication. Reliability estimates should be greater than .70 (Boone et al., 2014).

Rasch analysis produces a score for each participant (person measure), which indicates each participant's standing on a latent scale. Likewise, it also generates a score for each item, which indicates whether it is easy or difficult for participants to agree with each item (item difficulty). The distribution of person measures and item measures can be compared graphically using Wright maps produced in WINSTEPS to examine for floor and ceiling effects in dimensions of the PRI. These item-person maps depict the distribution of persons on the left and the distribution of items on the right. At the bottom of the continuum, fathers who reported lower levels of maternal gatekeeping are represented. Items that were relatively easier for fathers to endorse (i.e., occurred frequently) are also represented at the bottom. In contrast, the top part of the continuum represents fathers who report more frequent occurrences of maternal gatekeeping, as well as items that are harder or more challenging for fathers to endorse (i.e., higher difficulty level). Because items and persons are presented on the same logit scale, Wright maps clearly depict which items measure participants well. Item and person measured were rescaled from 0–100 to facilitate interpretability.

We assessed whether the rating scale model (RSM) or partial credit model (PCM) was most appropriate for analyzing data. Thresholds demarcate the probability boundaries between one response category and the adjacent response category. For example, a 6-point Likert scale has five thresholds that separate the six categories and indicate that there is an equal probability that the respondent selects a response category or an adjacent one. The RSM assumes that the thresholds for the categories are equivalent and identical for all items, whereas the PCM computes unique thresholds for each item (Boone et al., 2014). To determine whether the RSM or PCM should be used in the present analysis, a chi-squared difference test was conducted to examine whether the PCM significantly improved model fit. Results indicated the PCM did not significantly improve fit, so the RSM model was used. Finally, consistent with Rasch analysis standards, the data were "stacked" so that each participant had two sets of observations—one observation from 3-months and one observation from 9-months postpartum.

# Results

## Exploratory Factor Analyses

At 3-months postpartum, the initial parallel analysis recommended retaining four factors. However, Factor 4 had a trivial number of items (Items 31, 21, 10 and 35). Because factors with fewer than 5 items are considered unstable (Tabachnick & Fidell, 2007), these items were not retained. Parallel analyses were conducted a second time with the exclusion of these items and recommended retaining three factors. However, Items 26 and 19 had high cross-loadings (> .30) on two different factors and were eliminated, as items that load on more than one factor artificially raises inter-factor correlations and make it challenging to meaningfully interpret factors (Pett et al., 2003). Parallel analyses were conducted with the exclusion of these items and recommended retaining 3 factors. Next, parallel analyses were conducted at 9-months postpartum with the exclusion of items that did not function well at 3-months postpartum (Items 10, 19, 21, 26, 31, and 35). Three factors were recommended for retention. EFA results indicated that all factors had an appropriate number of items with sufficient factor loadings.

### Factor rotation.

A promax rotation of the gatekeeping items from the Parental Regulation Inventory at 3- and 9- months postpartum was conducted and, the reference structure matrix was then interpreted. To maintain consistent factor structures at both time points, only items that loaded on the same factors at 3- and 9-months postpartum, respectively, were retained. Next, factors were inspected to determine if items consistently reflected the underlying gatekeeping construct. Although EFA analyses are heavily data-driven, the final solution should be evaluated to ensure it is theoretically sound (Pett et al., 2003). Five items on Factor 1 (i.e. “Leave the house so you don’t have a choice,” “Refuse to do it herself,” “Tell you to do a child care task,” “Hint that work needs to be done,” and “Ask for help by talking through the baby.”) did not theoretically cohere with the other items loading on Factor 1 (gate closing), which relate to mothers’ use of negative verbal or nonverbal behaviours that are more characteristics of maternal gate closing. Mothers who “refuse to do [a task]” or “leave the house so [fathers] do not have a choice” might actually be taking an adverse approach to encouraging father involvement in childrearing. Thus, these items were eliminated from the gate closing factor (Items 11, 12, 9, 5, and 1). In total, the final gate closing factor included 9 items. Additionally, nine items loaded onto a second factor (gate opening) that related to mothers’ use of behaviours or verbalizations that encouraged father involvement in childrearing. Five items loaded on a third factor (communication) that related to mothers’ attempts to discuss, explain, or instruct fathers. At 3-months and 9-months postpartum the gate closing factor was negatively correlated with gate opening ($r_{3mo}$ = −.36; $r_{9mo}$=−.21) and positively correlated with communication ($r_{3mo}$ = 40; $r_{9mo}$=.28). The communication factor was positively correlated with maternal gate opening ($r_{3mo}$ = .05; $r_{9mo}$=.27) at 9-months (Table 1).

### Rasch Analysis for Gate Opening, Gate Closing, and Communication Factors

—One gate closing item (Item 34) and two gate opening items (Items 17 and 18) had MNSQ values slightly outside the acceptable range and were flagged. Extreme outlier responses

with z-scores greater than |3.00| on flagged items were set to missing. This approach reduces bias in person measures (Wright & Stone, 1979). In total, less than 1% of values were set to missing on gate closing and gate opening factors. Although our approach aligns with recommended practice (Gugiu et al., 2019; Linacre, 2002; Wright & Stone, 1979), sensitivity analyses were conducted to ensure there was minimal impact on person and item measures. Item measures were cross-plotted with and without setting misfitting responses to missing. All item measures fell within a 95% confidence interval regardless of the approach. Once outlier responses were addressed, Items 34 and 18 were still slightly misfitting (i.e., MNSQ = 1.6). To determine if retaining the items would bias person measures, sensitivity analyses were conducted by cross-plotting person measures for gate closing and gate open factors, respectively, with and without the flagged items. All person measures fell within a 95% confidence interval, which indicates including the flagged items does not bias person measures. Thus, items were retained. It is important to note that because the Rasch model can compute person measures that are free of the items used to measure them (Boone et al., 2014; Wright & Stone, 1979), setting outlier responses to missing does not bias person measures. Hence, the strategy of making data fit the Rasch model improves measurement. All items on the communication factor had a MNSQ value in the acceptable range.

Other indicators of model fit were also evaluated. Items with an expected point-measure correlation coefficient less than .30 are considered strong candidates for deletion (Persch et al., 2015), as a low correlation indicates that the item does not measure the latent construct very precisely. All gatekeeping items had point-measure correlations that were greater than .30. Person and item measures for each gatekeeping factor exceeded the .70 reliability standard. Finally, category probability curves recommended retaining 6 response categories, as each category was ordered as required and formed distinctive regions. Additionally, Andrich thresholds followed a monotonic progression, as expected.

### Wright maps.

Because participants and items are presented on the same logit scale, the researcher can draw conclusions about gaps along the continuum, redundant items, or floor and ceiling effects. As depicted on the annotated figures, floor effects occur when several items cluster at the top of the continuum and participants cluster toward the bottom (average item difficulty level is greater than average person measure). Ceiling effects occur when several participants cluster at the top of the continuum and items cluster near the bottom (average item difficulty level is less than average person measure).

Figure 1 presents the gate closing Wright map after item difficulty and person measures were rescaled to a 0 – 100 scale to facilitate interpretability, wherein low values designate less gate closing behavior and vice versa. At the centre of this figure is a vertical logit scale with the distribution of person measures on the left side and the distribution of item difficulty levels on the right side of the vertical line. Examination of this figure clearly illustrates that, although the instrument was capable of measuring fathers' perceptions of higher maternal gate closing behavior, items did not measure milder forms of this behavior. Item difficulty levels ranged from 45.36 to 58.08 logits. Person measures ranged from 61.75 to 1.67 logits. The mean item difficulty was 51.39 logits (i.e., extent to which items are,

on average, easy for participants to endorse), with a standard error of .65, while the mean person measure was 33.63 logits (i.e., extent to which, on average, participants embody latent trait), with a standard error of 5.13. This factor exhibited a floor effect, indicating several gate closing survey items were challenging for participating fathers to endorse.

Figure 2 presents the gate opening Wright map after scores were rescaled to a 0 – 100 scale. Item difficulty levels ranged from 44.64 to 57.53 logits. Person measures ranged from 25.50 to 92.83 logits. The mean item difficulty was 47.74 logits, with a standard error of .49, while the mean person measure was 49.95 logits, with a standard error of 3.15. On average, the survey difficulty level was very close to the average person measure, indicating the gate opening survey items were at an appropriate level for fathers. However, there floor and ceiling effects were observed.

Figure 3 presents the communication Wright map after scores were rescaled to a 0 – 100 scale. Item difficulty levels ranged from 47.01 to 48.27 logits. Person measures ranged from 52.22 to 41.98 logits. The mean item difficulty was 47.24 logits, with a standard error of .07, while the mean person measure was 47.24 logits, with a standard error of .55. On average, the survey difficulty level was very close to the average person measure, indicating the communication survey items were at an appropriate level for a father who reported average levels of maternal communication. However, there was evidence of floor and ceiling effects.

## Discussion

Strong measurement in family psychology research supports the advancement of sound theoretical models. Rasch analysis is a powerful technique often used to validate measures used in medicine and education (Boone et al., 2014). However, it is rarely applied to validate measures in family psychology. With the advancement of psychometric theory and research, Rasch analysis can be applied to better evaluate and improve survey instruments and conceptual models. This study served the joint purpose of 1) illustrating the utility of Rasch analysis techniques for family psychologists, and 2) applying these techniques to validate a measure of maternal gatekeeping: The Parental Regulation Inventory (PRI). Multiple theoretical perspectives on maternal gatekeeping have emerged in recent years (Fagan & Cherson, 2017; Puhlman & Pasley, 2013; Schoppe-Sullivan et al., 2008), yet advancements in maternal gatekeeping measurement have lagged behind. Thus, this study offers critical insight into the structure and validity of the PRI.

Prior to conducting Rasch analyses, the dimensionality of the PRI was assessed. Results indicated nine items loaded on a gate closing factor, nine loaded on a second gate opening factor, and five loaded on a third communication factor. This structure aligns with models that distinguish between encouragement (gate opening) and discouragement (gate closing) domains of maternal gatekeeping (Fagan & Cherson, 2017; Puhlman & Pasley, 2013; Schoppe-Sullivan et al., 2008). Additionally, results support the theory that maternal gatekeeping is multidimensional. In contrast to prior research that has considered maternal gatekeeping as a unidimensional construct—generally reflecting maternal beliefs concerning the importance of fathers (Fagan & Barnett, 2003)—researchers should be careful to distinguish between the multiple aspects of maternal gatekeeping. Beyond gate

opening and gate closing, a third factor focused on mothers' strategies for communicating her concerns or feelings toward fathers emerged. Communication has not been previously introduced as an aspect of maternal gatekeeping. However, Feinberg (2003) has considered communication an important part of the coparenting relationship. Parents may strive to manage family interactions by controlling patterns of communication with each other. Mothers' efforts to instruct and teach fathers might promote parents' joint responsibility to meet children's physical and emotional needs (Feinberg, 2003). Mothers' communication style is a potentially malleable aspect of maternal gatekeeping that could be targeted in parenting programmes. Although more research is needed to examine the implications of maternal communication for father involvement and child social-emotional adjustment, this study provides a first step in identifying communication items.

It is promising that items loading on gate closing and gate opening latent factors overlap conceptually with prior maternal gatekeeping research that has measured these domains (Schoppe-Sullivan et al., 2015). Additionally, reliability for the three dimensions of maternal gatekeeping was acceptable, ranging from .97 to .98. Person reliabilities were also high, ranging from .78 to .86. Further, point-measure correlations were all acceptable across the maternal gatekeeping dimensions. The Wright maps produced for each gatekeeping dimension offer insight into how well individual items measured participating fathers' perceptions of maternal gatekeeping. Recommendations for enhancing these scales are organized by gatekeeping dimension.

## Gate Closing Model

Inspection of the Wright map revealed several gate closing items were clustered at the top of the continuum, indicating the items might be challenging for fathers to endorse (i.e., several fathers reported lower occurrences of these types of behaviours). This pattern indicates a floor effect, in which the existing gate closing items target individuals who report high occurrences of gate closing well. To better target fathers who report lower levels of gate closing, researchers should apply maternal gatekeeping theory to identify gate closing behaviours that might occur more frequently in families—and thus be "easier" for fathers to endorse. Rescaled, the easiest item for participants to agree with was Question 20, which had an item difficulty level of 45.36. This question asked fathers to report how frequently mothers engaged in subtle gate closing behaviour: "When you do something that your baby's mother doesn't approve of regarding child care or with your baby, how often does she…show you that she is angry or irritated." Although this item had the lowest item difficulty level, it was still hard for fathers to report this behaviour occurred frequently. Thus, in order to better measure participants at lower levels of reported gate closing, easier items should be added in future administrations of the PRI survey. A less "difficult" measure may include more subtle behaviours that are more likely to occur as forms of maternal gate closing in families in dual-earner U.S. families. Existing qualitative work suggests mothers may limit father involvement in subtle ways by taking on tasks and making decisions with limited input from the father. An example of a subtle gate closing behaviour identified by fathers was painting the baby's room the color the mother wanted, limiting fathers' opportunities to pick out baby clothes, or putting on a different baby outfit after the father already dressed the baby (Hauser, 2012).

### Gate Opening Model

For maternal gate opening items, inspection of the Wright map indicated the average person measure score and item measure score were similar in magnitude. Those fathers with person measures near the middle of the distribution were measured relatively well. However, floor and ceiling effects were observed. In order to better measure participants at extreme scores along the continuum, easier as well as more challenging gate opening questions, should be added to the PRI. Rescaled, the easiest item for participants to agree with was Question 16, which had an item difficulty level of 44.64. In this item, fathers were asked to report the frequency with which mothers encouraged high quality father-child relationships by answering the question, "How often does your baby's mother do the following to encourage you to be involved in child care and with your baby…tell you how happy you make your baby?" Due to the floor effect, the difficulty level of this question was higher than several respondents' person measures. Future maternal gatekeeping research should incorporate questions that might be easier for new fathers to endorse. A more subtle behavior that mothers engage in to support fathers could be signaling enjoyment while watching the father interact with the baby (i.e., smiling or laughing).

Rescaled, the hardest item for participants to agree with was Question 18, which had a difficulty level of 57.53. In this item, fathers were asked to respond to the following question, "How often does your baby's mother do the following to encourage you to be involved in child care and with your baby…arrange activities for you and your child to do together." To better capture those participants who report very high levels of maternal gate opening, future research implementing the PRI should include several questions that would be more challenging for fathers to endorse. An example of additional items could be, "How often does your baby's mother ask you to teach your baby a new skill (i.e., clap)?" or "How often does your baby's mother ask you to help her pick out the baby's outfit?" Rasch analysis encourages maternal gatekeeping researchers to think more thoroughly about the range of gate opening behaviors mothers might exhibit in family interactions and can be used to assess whether item difficulty levels align with the researcher's expectations. Better targeting fathers who perceive very high and very low levels of maternal gate opening will better enable researchers to disentangle the implications of high and low levels of maternal gate opening for fathers' involvement in childrearing.

### Communication Model

Communication items are focused on the mothers' strategies for instructing or teaching fathers. Floor and ceiling effects were evident in the Wright map for communication. To better capture mothers who exhibit high or low levels of communication, both easier and more challenging communication items should be included in a future administration of the PRI. Rescaled, the easiest item for participants to agree with was Question 24, which had an item difficulty level of 47.01. In this item, fathers were asked to evaluate the following question, "When you do something that your baby's mother doesn't approve of regarding child care or with your baby, how often does she try to discuss her feelings about it with you?" The difficulty level of this question was higher than several respondents' average item measures across both time points. Thus, communication questions that are easier for new fathers to agree with should be incorporated in the PRI. For example, researchers could

include the following item, “How often does your baby’s mother tell you about something she knows the baby likes?”

Rescaled, the hardest item for participants to agree with was Question 22, which had a difficulty level of 48.27. In this question, fathers were asked, “When you do something that your baby’s mother doesn’t approve of regarding child care or with your baby, how often does she…tell you what she thinks you did wrong.” As a result of this ceiling effect, many participants had person measures that were much higher than individual item difficulty levels. To better capture those participants who report very high levels of maternal communication, future research implementing the PRI should include several questions that would be more challenging for fathers to endorse. An example of an item that could be included in a future administration might be, “How often does your baby’s mother give you step-by-step instructions for when you are in charge of childcare?” This item is an example of mothers’ communicating structured expectations for father involvement.

The detection of this factor highlights the importance of conceptualizing maternal gatekeeping as including multiple types of behaviors that occur in daily family interactions – in addition to gate closing and gate opening. It is important to note that, at 5 items, this factor only meets the minimum required number of items in a factor to be named (see Pett et al., 2003). Thus, we would issue caution in the use of this factor until further research is undertaken. Focus groups would help researchers identify a broader range of strategies mothers use to communicate their expectations to fathers.

## Conclusion

Rasch analysis provides a framework for evaluating the validity of surveys and identifying directions for survey improvement. Applied to the measurement of fathers’ perceptions of maternal gatekeeping, Rasch analysis enabled a detailed examination of the strengths and weaknesses of the Parental Regulation Inventory (PRI; Van Egeren, 2000). Future administrations of the PRI should include gate closing items that are easier for fathers to endorse, gate opening items that are both easier and more challenging for fathers to endorse, and communication items that are both easier and more challenging for fathers to endorse. The incorporation of additional items will allow for better measurement of fathers’ perceptions of maternal gatekeeping. Improved maternal gatekeeping measurement, in turn, will contribute to the development of a more unified maternal gatekeeping theory. Although we suggested several gatekeeping items that could be included in a revised version of the PRI, additional items could be developed by engaging fathers via qualitative research (i.e., focus groups) for discussions about the full range of maternal gatekeeping behaviours they observe. Across all dimensions of maternal gatekeeping, it appears gatekeeping behaviours may be even more subtle than previously assumed. Future questionnaire items about specific behaviours may need to be phrased more neutrally so that fathers are free to endorse or not endorse items.

This study is an important first step in defining the structure of maternal gatekeeping and examining the quality of individual items. However, the PRI is just one of a few maternal gatekeeping measures (i.e., Allen & Hawkins, 1999; Fagan & Barnett, 2003; Puhlman & Pasley, 2017). Researchers interested in administering a revised version of the PRI should

carefully review items or dimensions assessed by other measures to consider whether they could be included in the PRI. For example, a measure of maternal gatekeeping developed by Puhlman and Pasley (2017) included a control dimension, which refers to mothers' efforts to impose her will or manage fathers' interactions with the child (i.e., supervise him). The PRI does not include items that explicitly target maternal control. However, maternal gatekeeping theory would suggest that many gatekeeping behaviors are rooted in a power dynamic between parents (Allen & Hawkins, 1999).

Finally, it is important to note this study examined fathers' perceptions of maternal gatekeeping. How fathers perceive mothers' behaviours may be most important for the quantity and quality of fathers' subsequent involvement. Measures of gatekeeping behaviours should be analyzed together with fathers' more general perceptions of the coparenting relationship and characteristics of the fathers. Such an approach would enable researchers to better consider maternal gatekeeping within the family context. Asking fathers how they feel about certain types of maternal gatekeeping behaviours might help researchers identify the ways in which maternal gatekeeping is perceived by fathers.

This study is not without limitations. A key limitation is the demographic composition of our sample. The factor structure of the PRI was identified using an exploratory factor analysis of a sample of dual-earner, married or cohabiting fathers. This factor structure might not apply to fathers embedded in different family forms, for example, non-resident fathers. Future research should examine the performance of this measure in diverse family contexts. Additionally, the PRI is a single measure of maternal gatekeeping and might not capture other aspects of gatekeeping, such as mothers' beliefs about fathers' role in childrearing.

Notwithstanding these limitations, this study represents an important first step in examining the underlying factor structure of the PRI, as well as identifying the quality of individual gatekeeping items using Rasch analysis. This will inform future developments in maternal gatekeeping measurement, as well as provide an example for other family psychology researchers interested in implementing Rasch analysis. Rasch analysis encourages researchers to think more explicitly about the range of behaviors underlying the latent construct of interest and develop items that better align with behaviors reported in the population of interest. Wright maps are a tool unique to Rasch analysis and provide a visual display of the degree to which participants embody a particular latent trait, as well as the extent to which items represent participants. By using the many features of Rasch analysis, family psychologists can write better items, improve measurement of participants, and ultimately developed more refined conceptual models.

## Acknowledgments

This paper and its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NSF, NICHD, Pennsylvania State University—Shenango, or The Ohio State University. The New Parents Project was funded by the National Science Foundation (CAREER 0746548, Schoppe-Sullivan), with additional support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD; 1K01HD056238, Kamp Dush), and The Ohio State University's Institute for Population Research (NICHD R24HD058484) and program in Human Development and Family Science. We sincerely thank the many graduate and undergraduate students who recruited for, and collected, entered, and coded the data of the New Parents Project as well as the families who participated in the research.

We also acknowledge Claire M. Kamp Dush's invaluable contribution to the design and execution of the New Parents Project.

## References

Allen SM, & Hawkins AJ (1999). Maternal gatekeeping: Mothers' belifes and behaviors that inhibit greater father involvement in family work Journal of Marriage and Family, 61, 199–212. doi:10.2307/353894

Altenburger LE, Lang SN, Schoppe-Sullivan SJ, Kamp Dush, & Johnson S. (2017). Toddlers' differential susceptibility to the effects of coparenting on social-emotional adjustment. International Journal of Behavioral Development, 41(2), 228–237. doi:10.1177/0165025415620058 [PubMed: 31892763]

Altenburger LE, Schoppe-Sullivan SJ, & Dush CMK (2018). Associations between maternal gatekeeping and fathers' parenting quality. Journal of Child and Family Studies, 27(8), 2678–2689. doi:10.1007/s10826-018-1146-9.

Baghaei P. (2008). The Rasch model as a construct validation tool. Rasch Measurement Transactions, 22(1), 1145–1146.

Bond T, & Fox CM (2007). Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Earlbaum Associates, Inc. Publishers

Boone WJ, Staver JR, & Yale MS (2014). Rasch analysis in the human sciences. Dordrecht, The Netherlands: Springer.

Boone WJ, Townsend JS, & Staver J. (2011). Using Rasch theory to guide the practice of survey development and survey data analysis in sicience education and to informe science reform efforts: an exemplar utilizing STEBI self-efficacy data. Science Education, 95(2), 258–280. doi: 10.1002/sce.20413

Brown TA (2014). Confirmatory factor analysis for applied research. New York: Guilford Press.

Cabrera NJ, Fitzgerald HE, Bradley RH, & Roggman L. (2014). The ecology of father-child relationships: An expanded model. Journal of Family Theory & Review, 6(4), 336–354. doi:10.1111/jftr.12054

Cannon EA, Schoppe-Sullivan SJ, Mangelsdorf SC, Brown GL, & Sokolowski M. (2008). Parent characteristics as antecedents of maternal gatekeeping and fathering behavior. Family Process, 47(4), 501–519. doi:10.1111/j.1545-5300.2008.00268.x [PubMed: 19130790]

Coley RL, Lewin-Bizan S, & Carrano J. (2011). Does early paternal parenting promote low-income children's long-term cognitive skills? Journal of Family Issues, 32(11), 1522–1542. doi:10.1177/0192513X11402175

Cummings EM, Goeke-Morey MC, & Raymond J. (2004). Everyday marital conflict and child aggression. Journal of Abnormal Child Psychology, 32(2), 191–202. doi:10.1023/B:JACP.0000019770.13216.be [PubMed: 15164860]

Doherty WJ, Kouneski EF, & Erickson MF (1998). Responsible fathering: An overview and conceptual framework. Journal of Marriage and Family, 60(2), 277–292. doi:10.2307/353848

Duncan PW, Bode RK, Lai SM, Perera S, & Investigators, G. A. i. N. A. (2003). Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. Archives of Physical Medicine and Rehabilitation, 84(7), 950–963. doi:10.1016/S0003-9993(03)00035-2 [PubMed: 12881816]

Elliot R, Fox CM, Beltyukova SA, Stone GE, Gunderson J, & Zhang X. (2006). Deconstructing therapy outcome measurement with Rasch analysis of a measure of general clinical distress: The Symptom Checklist-90-Revised. Psychological Assessment, 18(4), 359–372. doi:10.1037/1040-3590.18.4.359 [PubMed: 17154757]

Elliott MC, Shuey EA, & Leventhal T. (2016). Adolescent functioning in housing and family contexts: A mixed methods study. Journal of Family Psychology, 30(6), 676–686. doi:10.1037/fam0000220 [PubMed: 27513286]

Fagan J, & Barnett M. (2003). The relationship between maternal gatekeeping, paternal competence, mother's attitudes about the father role, and father involvement. Journal of Family Issues, 38(5), 633–653. doi:10.1177/0192513X03256397

Fagan J, & Cherson M. (2017). Maternal gatekeeping: The associations among facilitation, encouragement, and low-income fathers' engagement with young children. Journal of Family Issues, 38(5), 633–653. doi:10.1177/1049909111402187

Fagan J, & Palkovitz R. (2011). Coparenting and relationship quality effects on father engagement: Variations by residence, romance. Journal of Marriage and Family, 73(3), 95–131. doi:10.1111/j.1741-3737.2011.00834.x

Feinberg ME (2003). The internal structure and ecological context of coparenting: A framework for research and intervention. Parenting: Science and Practice, 3(2), 95–131. doi:10.1207/S15327922PAR0302_01 [PubMed: 21980259]

Fowers BJ, Laurenceau J-P, Penfield RD, Cohen LM, Lang SF, Owenz MB, & Pasipanodya E. (2016). Enhancing relationship quality measurement: The development of the Relationship Flourishing Scale. Journal of Family Psychology, 30(8), 997–1007. doi:10.1037/fam0000263 [PubMed: 27918187]

Glorfeld LW (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. Educational and Psychological Measurement, 30(8), 377–393. doi:10.1177/0013164495055003002

Gordon RA (2015). Measuring constructs in family science: how can item response theory improve precision and validity? Journal of Marriage and Family, 77(1), 147–176. doi:10.1111/jomf.12157 [PubMed: 25663714]

Guadagnoli E, & Velicer WF (1988). Relation of sample size to the stability of component patterns. Psychological Bulletin, 103(2), 265–275. doi:10.1037/0033-2909.103.2.265 [PubMed: 3363047]

Gugiu C, Coryn CL, & Applegate B. (2010). Structure and measurement properties of the Patient Assessment of Chronic Illness Care Instrument. Journal of Evaluation in Clinical Practice, 16(3), 509–516. doi:10.1111/j.1365-2753.2009.01151.x [PubMed: 20210824]

Gugiu PC, Gugiu MR, Barnes M, Gimbert B, & Sanders M. (2019). The development and validation of the Parental Involvement Survey in their Children's Elementary Studies (PISCES). Journal of Child and Family Studies, 28(3), 627–641. doi:10.1007/s10826-018-1294-y

Hauser O. (2012). Pushing daddy away? A qualitative study of maternal gatekeeping. Qualitative Sociology Review, 8(1), 34–59.

Jensen TM, & Shafer K. (2013). Stepfamily functioning and closeness: Children's views on second marriages and stepfather relationships. Social Work, 58, 127–136. [PubMed: 23724576]

Linacre JM (2002). What do infit and outfit, mean-square and standardized mean. Rasch Measurement Transactions, 16(2), 878.

Linacre JM (2020). Winsteps® Rasch measurement computer program user's guide. Beaverton, Oregon: Winsteps.com.

Linacre JM, & Wright BD (1989). The "length" of a logit. Rasch Measurement Transactions, 3(2), 54–55.

Meteyer K, & Perry-Jenkins M. (2010). Father involvement among working-class, dual-earner couples. Fathering, 8(3), 379–403. doi:10.3149/fth.0803.379

Persch AC, Gugiu PC, Velozo CA, & Page SJ (2015). Rasch Analysis of the Wrist and Hang Fugl-Meyer: Dimensionality and Item-Level Characteristics Journal of Neurological Physical Therapy, 39(3), 185–192. doi:10.1097/NPT.0000000000000096

Pett MA, Lackey NR, & Sullivan JJ (2003). Making sense of factor analysis: The use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage Publications, Inc.

Phillipson S, & McFarland L. (2016). Australian parenting and adolescent boys' and girls' academic performance and mastery: The mediating effect of perceptions of parenting and sense of school membership. Journal of Child and Family Studies, 25(6), 2021–2033. doi:10.1007/s10826-016-0364-2

Puhlman DJ, & Pasley K. (2013). Rethinking maternal gatekeeping. Journal of Family Theory & Review, 5(3), 176–193. doi:10.1111/jftr.12016

Puhlman DJ, & Pasley K. (2017). The maternal gatekeeping scale: Constructing a measure. Family Relations, 66(5), 824–838. doi:10.1111/fare.12287

Scheffer J. (2002). Dealing with missing data. Research Letters in the Information and Mathematical Sciences, 3(1), 153–160.

Schoppe-Sullivan SJ, Altenburger LE, Lee MA, Bower DJ, & Kamp Dush CM (2015). Who are the gatekeepers? Predictors of maternal gatekeeping. Parenting, 15(3), 166–186. doi:10.1080/15295192.2015.1053321 [PubMed: 27366115]

Schoppe-Sullivan SJ, Brown GL, Cannon EA, Mangelsdorf SC, & Sokolowski MS (2008). Maternal gatekeeping, coparenting quality, and fathering behavior in families with infants. Journal of Family Psychology, 22(3), 389–398. doi:10.1037/0893-3200.22.3.389 [PubMed: 18540767]

Tabachnick BG, & Fidell LS (2007). Using multivariate statistics. Boston, MA: Pearson.

Van Egeren LA (2000). The Parental Regulation Inventory Unpublished manuscript. Michigan State University. East Lansing, MI.

Wright BD (1997). A history of social science measurement. Educational Measurement: Issues and Practice, 16(4), 33–45. doi:10.1111/j.1745-3992.1997.tb00606.x

Wright BD, & Stone MH (1979). Best Test Design. Rasch Measurement.

Zvara BJ, Schoppe-Sullivan SJ, & Kamp Dush CM (2013). Fathers' involvement in child health care: Associations with prenatal involvement, parents' beliefs, and maternal gatekeeping. Family Relations, 62(4), 649–661. doi:10.1111/fare.12023 [PubMed: 26405366]

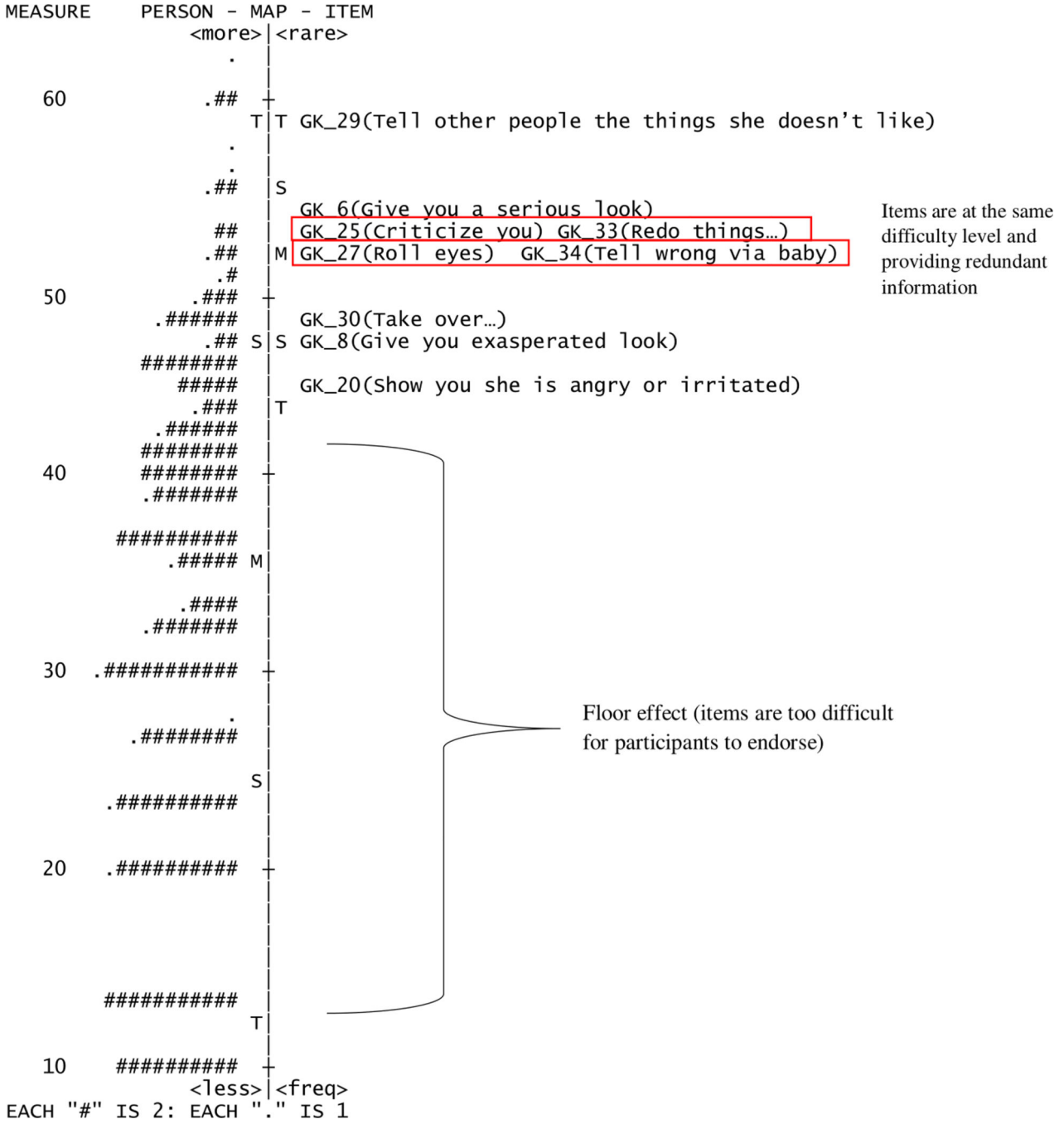


**Figure 1.**
Wright Map of item difficulty and person measures for gate closing items

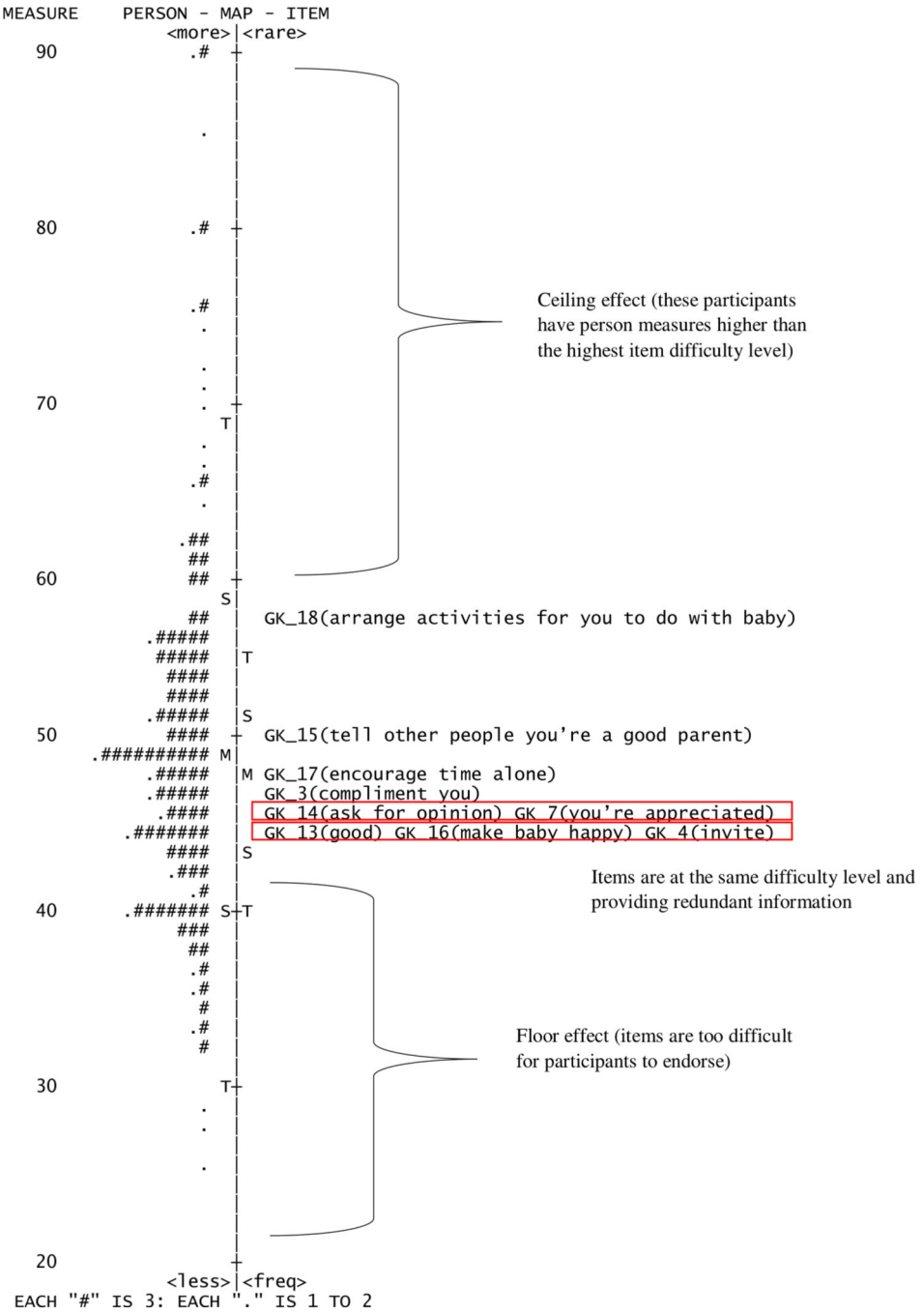


**Figure 2.**
Wright Map of item difficulty and person measures for gate opening items

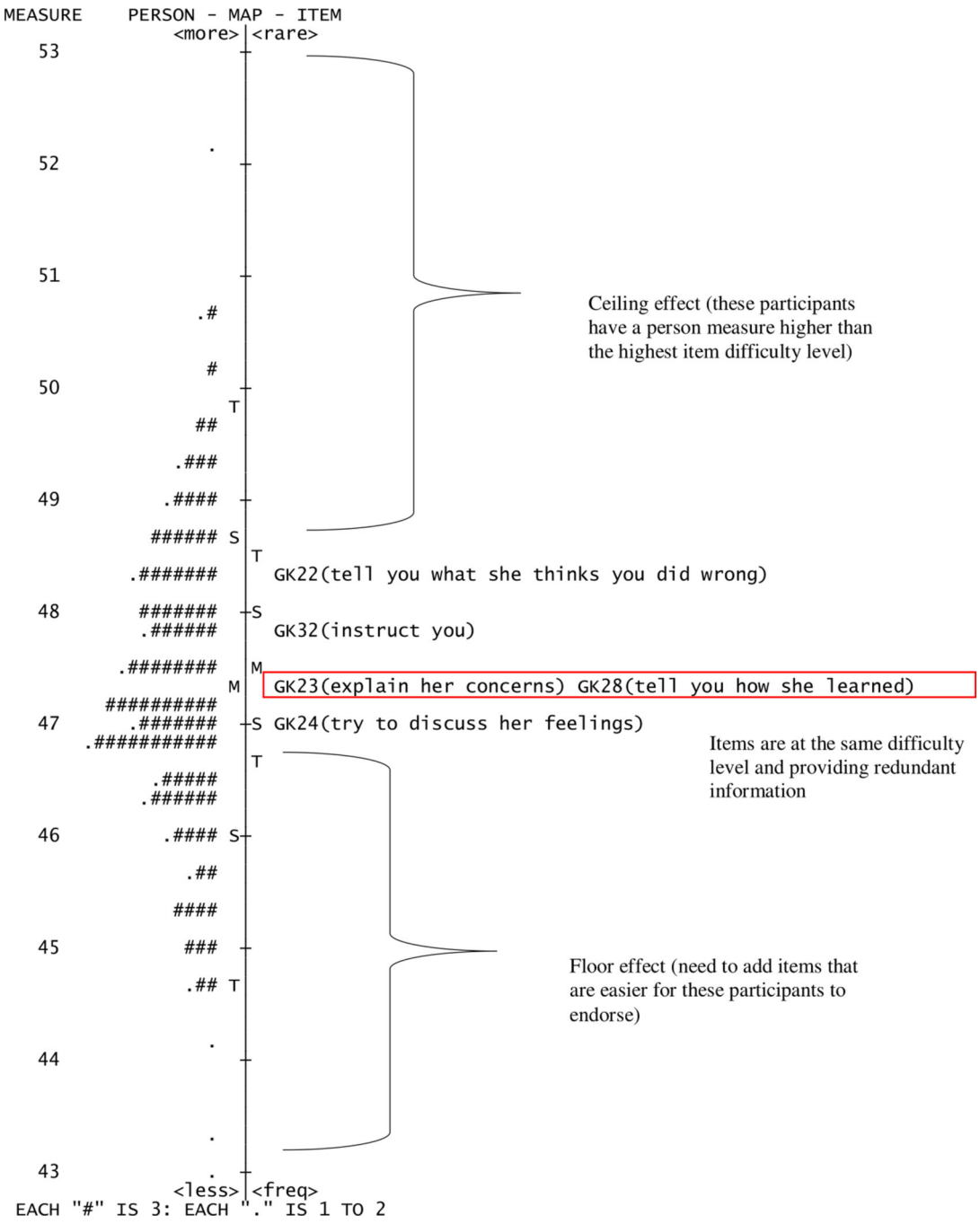


**Figure 3.**
Wright Map of item difficulty and person measures for communication items

## Table 1

Rotated factor loadings for maternal gatekeeping items at 3-months and 9-months postpartum

| Maternal Gatekeeping Items | Gate Closing | Gate Opening | Communication |
|---|---|---|---|
| 27. Look exasperated and roll her eyes | .79(.63) | | |
| 6. Give you a serious look | .63(.66) | | |
| 8. Give you an exasperated look | .59(.59) | | |
| 25. Criticize you | .56(.60) | | |
| 29. Tell other people the things she doesn't like | .55(.63) | | |
| 20. Show you that she is angry or irritated | .53(.63) | | |
| 33. Redo things after you are gone | .46(.63) | | |
| 30. Take over and do it her own way | .43(.64) | | |
| 34. Tell you what you did wrong by "talking through the baby" | .38(.43) | | |
| 13. Tell you what a good parent you are | | .80(.78) | |
| 3. Compliment you | | .76(.66) | |
| 16. Tell you how happy you make your baby | | .71(.68) | |
| 7. Let you know she appreciates your contributions | | .70(.67) | |
| 15. Tell other people about what a good parent you are | | .58(.73) | |
| 14. Ask for your opinion | | .53(.44) | |
| 4. Invite you to help | | .46(.47) | |
| 17. Encourage you to spend time alone with your baby | | .42(.32) | |
| 18. Arrange activities for you and your child to do together | | .33(.33) | |
| 23. Explain her concerns to you | | | .69(.71) |
| 28. Tell you how she has learned to handle similar situations | | | .56(.47) |
| 24. Try to discuss her feelings about it with you | | | .56(.50) |
| 22. Tell you what she thinks you did wrong | | | .51(.50) |
| 32. Instruct you | | | .50(.42) |

*Note*. Only factor loadings above .30 are shown. Factor loadings from 9-months postpartum are indicated in parentheses.