Feeling Appreciated Buffers against the Negative Effects of Unequal Division of Household Labor on Relationship Satisfaction

Amie M. Gordon[1]*, Emily Cross[2], Esra Ascigil[1], Rhonda Balzarini[3,4], Anna Luerssen[5], Amy Muise[2]

[1]Department of Psychology, University of Michigan; Ann Arbor, MI, 48109.

[2]Department of Psychology, York University; Toronto, Canada, M3J 1P3.

[3]Department of Psychology, Texas State University; San Marcos, TX, 78666.

[4]The Kinsey Institute, Indiana University; Bloomington, IN, 47405.

[5]Department of Psychology, Lehman College, City University of New York; New York

City, NY, 10468.

In Press, Psychological Science

Supplementary Tables and Figures

*Corresponding author: amiemg@umich.edu

Abstract

Decades of research from across the globe highlight unequal and unfair division of household labor as a key factor that leads to relationship distress and demise. But does it have to? Testing a priori predictions across three samples of individuals cohabiting with a romantic partner during the COVID-19 pandemic ($N = 2{,}193$, including 476 couples), we find an important exception to this rule. People who reported doing more of the household labor and perceived the division as more unfair were less satisfied across the early weeks and ensuing months of the pandemic, but these negative effects disappeared when people felt appreciated by their partners. Feeling appreciated also appeared to buffer against the negative effects of doing *less*, suggesting that feeling appreciated may offset the relational costs of unequal division of labor, regardless of who contributes more. These findings generalized across gender, employment status, age, socioeconomic status, and relationship length.

Keywords: Interpersonal Relationships, Relationship Satisfaction, Gratitude, Division of Labor

Statement of Relevance

The quality and stability of romantic relationships are among the strongest predictors of individuals', couples', and families' overall health and well-being. Hence, romantic relationships are a critical focus of researchers across disciplines. Unequal and unfair division of household labor are key factors that predict relationship distress and demise, and household chores are a top cited source of conflict in couples. While an equal and fair distribution of labor between partners is ideal, inequality is common: more than half of our participants reported their division of labor was uneven or unfair. Our findings from samples in the USA and Canada provide a new way of thinking about division of labor, revealing that the costs of bearing more of the household burden may only exist when people feel taken for granted by their partner. In contrast, people with more appreciative partners maintain satisfaction even when the division of household labor is unequal.

Feeling Appreciated Buffers against the Negative Effects of Unequal Division of Household Labor on Relationship Satisfaction

The quality and stability of romantic relationships are among the strongest predictors of overall health and well-being (Holt-Lunstad et al., 2010; Kansky, 2018; van Eldik et al., 2020), making romantic relationships a critical focus across disciplines. Decades of studies from across the globe highlight unequal and unfair division of household labor as a key factor that leads to relationship distress and demise (e.g., Adams, 1965; Qian & Sayer, 2016; Shockley & Shen, 2016; Thielemans et al., 2020; Van Yperen & Buunk, 1990). Deciding who makes dinner or pays the bills is a top source of conflict for couples (Jackson et al., 2016) with far-reaching consequences. For example, in a sample of over 3,000 Danish couples, Thielemans and colleagues (2020) found that couples with the most unequal division of labor were the most at risk for relationship dissolution. But are people always less satisfied when they do more around the house? We propose an important exception to this rule: doing more might not have the same negative effects when people generally feel appreciated by their partners.

Feeling appreciated—feeling recognized and valued, rather than taken for granted—helps couples maintain high quality relationships (e.g., Algoe, 2012; Gordon et al., 2012; ter Kuile et al., 2017), buffers against relationship insecurities (Park et al., 2019) and negative conflict patterns (Barton et al., 2015) and has been identified as an important factor in caregiving contexts (Amaro, 2017). Appreciation is also associated with how people feel about their household contributions. Feeling appreciated for doing chores can help people reframe the chores from a cost to a benefit (Berger & Janoff-Bulman, 2006) and people who perceive the division of labor as more equal and fair feel more recognized for their contributions (Blair & Johnson, 1992; Hawkins et al., 1995;

Klumb et al., 2006; Mikula et al., 2009). Here, we approach the association between appreciation and division of labor from a new angle. We examine a priori predictions about the buffering effects of feeling appreciated on concurrent relationship satisfaction and changes in relationship satisfaction over time when people *do not* see household labor as equally divided. We also examine satisfaction with the division of labor itself, as well as expectations of future relationship satisfaction, which is a strong predictor of future relationship outcomes, such as divorce (Baker et al., 2017).

We gathered our data during the COVID-19 pandemic, a period when couples spent more time at home with less outside help, often caring for children, and experienced significant changes in employment, including working from home. These lifestyle changes resulted in increased domestic labor as well as shifts in how labor was divided between partners (Craig & Churchill, 2021; Shafer et al., 2020; Shockley et al., 2020; Waddell et al., 2021), making the pandemic an especially relevant context in which to examine division of labor and the potential buffering effects of feeling appreciated.

The Current Research

At the beginning of the pandemic, we assessed division of labor, appreciation, and relationship satisfaction in three different samples as part of larger pandemic-related projects. In two of the three samples, participants completed four follow-up assessments spanning 9 and 6 months (Samples B & C).

In addition to testing our main prediction that feeling appreciated would buffer against the relationship costs of doing more of the household labor, we also tested several additional questions and alternative explanations. First, perceived unfairness plays an important role in how division of labor affects relationships (Adams, 1965; Shockley &

Shen, 2016). Even when doing more household labor, people may perceive it as fair if their partner does more in other areas (e.g., paid labor), or if it aligns with their ideology (e.g., traditional gender role beliefs; Lavee & Katz, 2002). Yet, when the division *is* perceived as unfair, inequity is particularly detrimental, undermining individual and relationship well-being (Adams, 1965; Shockley & Shen, 2016). Thus, we tested whether feeling appreciated would attenuate the costs of contributing more even when those contributions were perceived as unfair.

Second, evidence of the gendered nature of household labor is robust: women do more in mixed-gender couples, even among dual-earning couples (Shockley & Shen, 2016). Indeed, work on division of labor frequently focuses on gender. Accordingly, we examined gender differences in our primarily mixed-gender samples, including whether any buffering effects of feeling appreciated differed by gender. Consistent with past work, we expected women to report doing more household labor. However, we did not expect that the benefits of feeling appreciated would differ by gender.

Third, household labor receives much of the attention in the literature on division of labor (Shockley & Shen, 2016), but couples must also divide up other types of labor, including paid labor and childcare. We assessed these additional areas of labor in two of our samples (Sample A & B) to test whether the predicted buffering effects of feeling appreciated were consistent across different domains, or specific to household labor. We also considered the average across labor domains given that people may contribute less to household labor if they are contributing more in other domains. Although, the main effects of division of labor on relationship satisfaction might differ by domain (Shockley & Shen, 2016), we had no theoretical reason to expect differences in the buffering effect

of feeling appreciated. Additionally, given that household contributions may differ depending on employment status, we included employment status as a covariate and moderator in our analyses. Finally, we tested whether our results differed depending on several relevant factors, including socioeconomic status, age, and relationship duration, as well as whether the predicted effect held controlling for physical and mental health, critical variables likely to affect both division of labor and feelings about one's relationship during the pandemic.

Methods

Sample A Overview

Sample A includes data from 1,195 participants (including 135 couples) in the United States and Canada who completed a single assessment between April 20th, 2020 and February 12th, 2021.

Sample A Methods

Participant demographics for all three samples are shown in Table S1. For Sample A, we aimed for 1,000 participants and at least 100 couples based on feasibility and a priori power analyses conducted for Sample B (see Sample B methods). Data collection is ongoing, but we report on analyses from 1,195 participants who met inclusion criteria (i.e., completed at least 60% of the surveys and reported paying attention); 925 participated in the survey individually and 270 participated with a romantic partner (135 couples). Participants were recruited using ads on social media (Facebook, Twitter, Craigslist, Reddit). The study was open to individuals in the U.S. and Canada who were at least 18 years old, in a romantic relationship, and cohabiting with their partner during the pandemic. Although they were not required to do so, we encouraged participants to

invite their partner to complete the survey as well. Those who were willing received a link and a dyad ID to share with their partner. Participants were volunteers and did not receive compensation for their participation. Participants were informed that initial results would be posted online at the end of the first month of data collection. The study was reviewed by the University of Michigan Institutional Review Board (IRB) and was found to be exempt from IRB approval.

Sample A Measures

Table S2 displays descriptive statistics for all primary measures, and Table S4 displays correlations between key variables. We assessed ***feeling appreciated*** with three items from the Appreciation in Relationships Scale (AIR Scale; Gordon et al., 2012): "My partner often takes me for granted" (reverse scored), "My partner makes sure I feel appreciated," and "When I am with my partner, sometimes they will look at me excitedly and tell me how much they appreciate me". These items were selected as the consistently highest-loading items in factor analyses from multiple datasets using the full scale. The items were rated from 1 = *Completely disagree* to 7 = *Completely agree* and were averaged to create a mean score, with higher scores indicating feeling more appreciated (α = .86). ***Perceived division of labor*** was assessed with the question: ***"***How has the division of labor been handled since the pandemic began?" This was asked for three separate domains: household tasks, child-related tasks (when relevant), and financial contributions (-2 = *All me*, -1 = *More me*, 0 = *Equal*, 1 = *More my partner*, 2 = *All my partner*). Analyses were conducted on the separate items as well as their ***average (***all three items if people reported child-related tasks, otherwise the average of household and financial). *Relationship Satisfaction* was assessed in two ways: ***Relationship satisfaction***

over the prior week, "In the past week, how satisfied have you been with the following… Your relationship overall?" (1 = *Not at all satisfied* to 5 = *Completely satisfied*) and *expected changes in relationship satisfaction,* "Assuming that the pandemic is not resolved quickly and your current situation continues, how do expect your feelings and experiences will change in the weeks and months ahead relative to how you feel right now? …Your overall relationship satisfaction" (1 = *Less/Lower*, 5 = *No change*, 10 = *More/Higher*). We also assessed **satisfaction with division of labor** for each domain by providing participants with the question: "How satisfied are you with the division labor since the pandemic began?" and then listing each domain below the question with a separate Likert scale (1 = *Not at all satisfied* to 5 = *Completely satisfied*). Additional **demographic moderators** included gender (*Man, Woman, another identity*), employment status (*Not working, Employed Part-Time, Employed Full-Time*), household income, education, age, and relationship length (see Table S1). People who selected "Not sure" or "Would rather not say" for income and "Other" for education (< 5% of data) were not included in moderation analyses. **Physical health** was assessed with a single item tapping into people's sleep quality, "During the past week, how would you rate the overall quality of your sleep?" on a 4-point scale (1 = Very poor, 2 = Fairly poor, 3 = Fairly good, 4 = Very good). Our other pre-registered physical health variable, exercise, was not associated with relationship satisfaction. **Mental health** was computed using the Ultra-brief Screening Questionnaire for Depression and Anxiety (α = .87, Kroenke et al., 2009); 1 = *Not at all* to 4 = *Nearly every day*) which includes the following four items: "Feeling nervous, anxious, or on-edge," "Not being able to stop or control worrying," "Feeling

down, depressed, or hopeless," and "Little interest or pleasure in doing things." Items were reverse scored so higher scores indicate greater mental health.

Sample B Overview

Sample B includes data from 618 participants (including 151 couples) who completed a baseline survey (T1) in April 2020 as well as follow-up surveys in May 2020 (T2 $N = 556$), August 2020 (T3, $N = 432$), November 2020 (T4, $N = 362$), and February 2021 (T5, $N = 292$).

Sample B Methods

Participant demographics are shown in Table S1. Sample size was based on a priori power analyses. To capture small to moderate effects with dyadic data using the Actor-Partner Interdependence Model, we aimed to collect data from 150 couples (based on power analyses using the APIMPowerR shiny app; Ackerman & Kenny, 2016). At T1, a total of 618 participants met inclusion criteria (i.e., met eligibility criteria, completed at least 60% of the survey, reported paying attention, and provided coherent answers to open-ended questions); 316 participants completed the study individually and 151 completed it with a partner (i.e., 302 individuals). Using Prolific.co, people who were at least 18 years old, living in the U.S. and sheltering-in-place with their romantic partner at the time of recruitment (i.e., not working outside the home or leaving home except for essential business and exercise) were invited to complete the study. Participants were paid $3 for completing the T1 survey and $1.50-$2 for each additional survey. Participants who completed the first four surveys received a $1.50 bonus. Similar to Sample A, we asked participants if their partner would be willing to participate in the study as well, and if so, participants provided their partner's Prolific ID (if available) or

received a link to share the survey with their partner. The study was reviewed by the University of Michigan IRB and was found to be exempt from IRB approval.

<u>**Sample B Measures**</u>

Baseline. At Time 1 (T1), we used the same measures from Sample A to assess *feeling appreciated*, *perceived division of labor*, *relationship satisfaction over the prior week, expected changes in relationship satisfaction,* and *satisfaction with division of labor* (see Table S2 for descriptive statistics and Table S4 for correlations for key variables). We also assessed *demographic covariates/moderators* (gender, employment status, income, education, age, and relationship length), and *physical* and *mental health* ($\alpha = .87$) using the same variables as with Sample A. Descriptives for demographics are in Table S1.

Follow-ups. Descriptives for the key follow-up variables are in Table S3. In the follow up surveys at T2-T5, we assessed the extent to which people *felt appreciated by their partners over the prior week* with the item: "In the past week, how much have you felt appreciated by your partner?" (1 = *Not at all* to 5 = *Extremely*). At T4 and T5, we also asked about the extent to which they *felt appreciated for the contributions they made to their household* with the item: "How appreciated do you feel by your partner for the contributions you made to your household in the past week (financial, household tasks, child-related tasks, planning, etc)?" (1 = *Not at all*, 2 = *A little*, 3 = *Moderately*, 4 = *A lot*, 5 = *Extremely*). *Division of labor* and *satisfaction with division of labor* were assessed with the same items from Sample A, but were reframed to be about division of labor over the prior week. *Relationship satisfaction over the prior week* and *expected changes in relationship satisfaction* were assessed with the same items from baseline.

Sample C Overview

Sample C includes data from 380 participants (190 couples) who completed a baseline survey (T1) between April-June 2020. Participants were invited to complete three weekly surveys (W1-W3; Ns = 324-345) beginning one week after their baseline survey and a follow-up survey in November 2020 (T2, 4-6 months after T1, N = 293).

Sample C Methods

Participant demographics are shown in Table S1. Couples in Sample C were recruited through online advertisements (Craigslist, Kijiji, Facebook/Instagram) and research platforms (Honeybee, Research Stream). We aimed to recruit at least 100 couples based on feasibility and guidelines for dyadic analyses and recruited as many as possible through June 2020. Of the 196 couples (N = 392) who participated, six participants were removed for not passing attention checks or completing the relevant measures. Our final sample included 190 couples (N = 380). Eligibility criteria included that both partners agreed to participate, were at least 18 years old, lived together, were in a relationship for at least 6 months, had access to a computer and internet, and lived in the USA or Canada. Interested couples contacted the research team via email and were sent a brief eligibility survey, after which eligibility and interest were confirmed via voicemail message or a phone call with both partners. If both partners were eligible and agreed to participate, they were sent the link to complete a baseline survey. Then, each week for the next three weeks, participants were sent a survey. The weekly surveys did not include all baseline measures and some measures were truncated to reduce participant fatigue, increase efficiency, and minimize participant attrition. Participants were compensated $15 CAD ($12 USD) for completing the baseline survey, $5 CAD ($4 USD) for each weekly

survey, and $10 CAD ($8 USD) for completing a follow-up survey four to six months

after the baseline survey. The study was reviewed and approved by the York University's

Research Ethics Board (#e2020-109).

Sample C Measures

Baseline. At Time 1 (T1) we assessed ***feeling appreciated*** with two items from

the AIR Scale (Gordon et al., 2012): "My partner appreciates me" and "At times my

partner takes me for granted" [Reverse Scored]) rated from 1 = *Strongly disagree* to 7 =

Strongly agree. Items were averaged ($r = .49$, $p < .001$) to create a mean score, higher

scores indicate feeling more appreciated. Perceived ***household division of labor*** was

assessed by asking participants "Since the COVID outbreak, how have you and your

partner divided up household chores?" rated on the following scale: 1 = *I have done all of*

the housework, 2 = *I have done most of the housework,* 3 = *We have split up household*

chores equally, 4 = *My partner has done most of the housework,* 5 = *My partner has*

done all of the housework. Perceived ***fairness of division of labor*** was assessed by asking

participants "Since the COVID outbreak, to what extent do you think you and your

partner's division of household chores is fair?" rated on the following scale: 1 = *My*

partner has much more than their fair share, 2 = *My partner has done a bit more than*

their fair share, 3 = *It has been fair/evenly divided,* 4 = *I have done a bit more than my*

fair share, 5 = *I have done a lot more than my fair share*. In order to match the direction

of division of labor, items were reverse coded so that higher scores represent greater

unfairness to the partner. ***Relationship satisfaction*** was assessed with three items from

the Investment Model Scale (Rusbult et al., 1998): "I feel satisfied with our relationship."

"My relationship is much better than others' relationships." "Our relationship makes me

very happy" (α = .91: 1 = *Strongly disagree* to 7 = *Strongly agree*). Items were averaged to create a mean score with higher scores indicating higher relationship satisfaction. ***Physical health*** was captured via the item "How would you rate your physical health?" (1 = *poor*, 5 = *excellent*). ***Mental health*** was assessed using two items from the Ultra-brief Screening Questionnaire for Depression (Kroenke et al., 2009); "Since the outbreak of COVID, I have been feeling down, depressed, or hopeless", and "Since the outbreak of COVID, I have little interest or pleasure in doing things" (0 = *not at all*, 4 = *nearly every day*). Items were averaged to create a mean score (r = .71, p <.001).

Weekly. Each week perceived ***household division of labor*** was assessed by asking participants "In the last week, how have you and your partner divided up household chores?" and ***fairness of division of labor*** was assessed by asking participants "In the last week, to what extent do you think you and your partner's division of household chores is fair?" using the same response options as at baseline. ***Relationship satisfaction*** was assessed with one item: "In the last week, I feel satisfied with our relationship" (1 = *Strongly disagree* to 7 = *Strongly agree*).

Follow-up. At Time 2 (T2), participants completed the same measures as baseline as well as the same item from Sample B assessing ***feeling appreciated for household contributions***.

Analytic Strategy

Power considerations. Given that our sample sizes were based on feasibility constraints and power analyses for different effects (i.e., APIM) and that our primary analyses included interactions which tend to be smaller effects, we conducted sensitivity analyses using G*Power 3.1 to determine the smallest effects we were powered to detect

at 80% power with a two-tailed test. In order to account for the non-independence of our couple members, for each sample we calculated the *effective* sample size using the design effect equation ($N_{effective} = N/[1+(n_{cluster} - 1)*\rho]$ where N = total sample, $n_{cluster}$ = cluster size (in our case, 2) and ρ = intraclass correlation (ICC; i.e., measure of non-independence between couple members). When we had multiple outcomes, we used the largest ICC to provide the most conservative estimate. We then added the number of individual participants to the effective sample size for our couples and used that total sample size to run sensitivity analyses. For Sample A, our smallest effective sample size was 1,095, which was powered to detect effects as small as $\rho = .08$ (where ρ = population correlation coefficient). For Sample B, we focused on the baseline data given that additional time points would increase power. Our smallest effective sample size was 527, which was powered to detect effects as small as $\rho = .12$. For Sample C, we assessed sensitivity for both baseline and the 6-month follow-up. Our effective sample sizes were 258 (baseline) and 200 (follow-up), which were powered to detect effects as small as $\rho = .17$ and $.20$. Overall, these samples were sensitive enough to detect small effects.

     Analytic models. Our predictions and analytic strategies, including our linear and curvilinear associations between division of labor and relationship satisfaction and several additional supplementary analyses were posted online post-data collection but prior to analysis. Sample A was not part of the original preregistration, but we used the same approach as for our preregistration of Sample B baseline data (Preregistration for Sample B, Preregistration for Sample C). Additional analyses were conducted post-registration based on feedback (e.g., piecewise regressions in Figure 2, simple slopes in Table S12).

Given that all of our samples included participants nested within dyads, we used mixed-effects models to address the non-independence in our data. Models were run in SPSS Mixed Models (SPSS 27.0) and the nlme package in R (Pinheiro et al., 2017). This approach handles missing data, so for samples with both individuals and dyads, participants who did not have a partner were included in the models.

For the dyadic longitudinal data (i.e., weekly and monthly over-time analyses) in Samples B and C, we employed two-level cross-classified models to account for non-independence of partners within dyads and within time points. In initial models of Sample B which had all variables measured at multiple time points, we looked at time as a predictor of our main variables and found small effects (changes of less than .05 from time to time) for concurrent relationship satisfaction (positive) and expected relationship satisfaction (negative). However, time was not a significant predictor of either division of labor or feeling appreciated (ps > .3), suggesting time was not a confounding variable. We modeled separate random intercepts and slopes for each partner within the dyad, but treated the partners as indistinguishable and utilized compound symmetry matrices for the random effects to constrain the two partners to have the same parameters. Random slopes were modeled for time-varying predictors, but covariances between random effects were not modeled. When models failed to converge or random variances were unable to be computed, we removed those random slopes. The fixed effects estimates changed negligibly between models with and without random slopes. Figures represent results from models using only random intercepts due to the inability to place equality constraints on the random effects in R.

For analyses of concurrent and expected relationship satisfaction, we predicted

relationship satisfaction at the same time point. In order to analyze change in relationship

satisfaction over time, we created a change score (i.e., subtracting initial satisfaction from

follow-up satisfaction). To account for the fact that initial levels of satisfaction may be

related to changes in satisfaction, we also controlled for initial levels of satisfaction (i.e.,

relationship satisfaction measured at the same time point as our predictor variables). We

predicted change in satisfaction across 3 months in Sample B using the four follow-up

surveys, and from baseline to the weekly surveys and the 6-month follow-up in Sample

C.

We tested whether our effects held when including a binary gender code (woman

vs man, given this has been the focus in the literature) and employment status (working

vs not working), and tested gender and employment status as moderators, along with

income, education, age, and relationship length. To ensure sufficient power, we assessed

these additional moderators using the combined baseline data from Samples A & B

(which included the same measures). This provided us with a combined sample of 1,813

participants. Adjusting for the dependency between couple members left us with an

effective N of 1,617, which was sufficient to detect effects as small as $\rho = .07$.

Given prior research and theory (Adams, 1965; Thielemans et al., 2020; Van

Willigen & Drentea, 2001) suggest equality is ideal and inequality in either direction

(doing more *or* less) can be problematic, we preregistered that we would examine the

data for curvilinear trends. For any curvilinear associations between household division

of labor and relationship satisfaction, we tested whether there were differences in how

inequality was associated with relationship satisfaction when people reported they were

doing more versus their partner was doing more. To do this, we ran piecewise regressions testing the simple linear slopes from (1) "all me" to "equal" and (2) "equal" to "all my partner."

Thus, for our main hypotheses, *we tested up to four nested models with the following predictors:* (1) linear division of labor (DoL); (2) curvilinear DoL; (3) linear/curvilinear DoL + feeling appreciated moderating linear DoL; and (4) linear/curvilinear DoL + feeling appreciated moderating curvilinear DoL.

Relevant data, code, and materials are available online (Samples A & B, C).

Results

Division of labor (DoL) for each sample is shown in Figure 1. The most common response across samples was that both partners contributed equally to the division of labor (37% - 52% across samples), but a substantial number of participants reported unequal contributions, with many contributing more than their partners (28% - 47%). Descriptive statistics and correlations for the key variables in each sample are in Tables S2-S5. Overall, consensus between partners in our dyadic subset was high regarding household DoL (rs = 0.66-0.72).

Are people less satisfied when they feel they do more of the household labor?

Perceived household DoL had both significant linear and curvilinear associations with relationship satisfaction. In line with equity theory (Adams, 1965), people were most satisfied with an equal division of labor (Figure 2; Tables S6-S9). However, we found that the curvilinear effect was not symmetrical. Instead, whereas people tended to be less satisfied when they reported doing more, there were not consistent reductions in satisfaction when people reported their partner doing more (see Figure 2). This pattern

emerged for concurrent satisfaction (Samples A-C) as well as expected (Samples A & B) and actual (Samples B & C) changes in relationship satisfaction over time. Similar effects emerged when predicting satisfaction with household DoL (Table S10).

Does feeling appreciated buffer against the cost of doing more of the household labor? In all three samples, feeling appreciated moderated the association between perceived household DoL and concurrent relationship satisfaction. That is, people who felt more appreciated (+1 *SD*) tended to be equally satisfied regardless of their perceived household DoL, whereas those who felt less appreciated (-1 *SD*) were significantly less satisfied when they reported contributing more (Figure 3, rows 1-2; see Table 1 for simple slopes; full models in Tables S6-S7). This buffering effect was often present for both the linear and curvilinear effects of DoL.

The same effect emerged and was even stronger when assessing *changes* in relationship satisfaction across the early weeks and ensuing months of the pandemic (Samples B & C; Table S9). When people reported doing more, those who felt less appreciated became less satisfied over time whereas those who felt more appreciated maintained their levels of satisfaction (Figure 3, row 4). Effects for expected changes in satisfaction were less consistent (Samples A & B; Figure 3, row 3, Table S8), suggesting people may not realize how these factors affect their future satisfaction. We also found similar buffering effects when looking at satisfaction with DoL specifically (Samples A & B; Figure S1, Table S10) and feeling appreciated specifically for one's household contributions (Samples B & C; Table S11).

Although our hypotheses focused on the role of feeling appreciated when people reported doing more of the household labor, our interaction effects shown in Figure 3 and

the simple slopes in Table 1 suggest that feeling appreciated may also play a buffering role when people report their *partner* doing more [see Table S12 for simple slopes of feeling appreciated predicting satisfaction at differing levels of household DoL (i.e., I do more, Equal, Partner does more)]. Although feeling appreciated was generally a strong predictor of relationship satisfaction at all levels of DoL, the slopes were steeper when DoL was unequal. In other words, feeling appreciated may act as a buffer when division of labor is perceived to be unequal, regardless of who is perceived to be contributing more. We do note that the confidence intervals are larger when looking at reports of partners doing more (a less frequent occurrence), leading us to interpret these findings more cautiously.

Is feeling appreciated by one's partner beneficial even when household division of labor is perceived as unfair? Feeling appreciated by one's partner also buffered against the negative effects of perceived unfairness of DoL on relationship satisfaction concurrently and over time (Figure S2 & Table S13). In other words, feeling appreciated helped reduce the relational costs of feeling that one is doing more than one's fair share of the household labor.

Are these effects gendered? Consistent with prior work, across all of our samples, people tended to report women doing more of the household labor (Figure 1). However, including gender in our models did not meaningfully change our key effects (Table S14-S16), nor did gender moderate the DoL-concurrent relationship satisfaction link (tested in the combined baseline sample; Table S17). Thus, although women (versus men) did more of the household labor, we did not find evidence that gender differences meaningfully

explained or changed the buffering effects of appreciation in our primarily mixed-gender couples.

Do the same patterns emerge for perceived childcare and financial DoL?

Associations between perceived childcare and financial DoL and relationship satisfaction were weak or nonexistent, and feeling appreciated did not moderate the effects (ps > 0.29; Figure S3 & Table S18). Instead, there was only a main effect for feeling appreciated. The pattern for average DoL, which accounted for contributions across domains (e.g., doing less housework but more paid labor) appeared similar, albeit weaker, than household DoL. These results suggest household DoL may function differently than other labor domains when it comes to influencing the quality of relationships.

How robust are these effects? Using the combined baseline sample for power, we examined whether the buffering effects of feeling appreciated differed depending on employment status, income, education, age, or relationship length. Although people who were employed reported doing less household labor (rs = 0.21-0.25, Table S4), employment status did not explain or moderate our effects (Tables S19 & S20). Effects also did not differ by income, education, age, or relationship duration (Tables S21-S22) and the same pattern of effects emerged when accounting for mental and physical health at baseline (Tables S23 & S24).

Discussion

Division of household labor is a leading source of conflict for couples and an unequal or unfair division of labor between partners is associated with relationship dissatisfaction and dissolution. The COVID-19 pandemic and associated lockdowns

made this salient as many couples found themselves at home juggling work, childcare, and household demands (e.g., Craig & Churchill, 2021; Shafer et al., 2020; Shockley et al., 2020; Waddell et al., 2021). In the current work, we tested whether feeling appreciated by one's partner buffered against the relationship costs associated with doing more of the household labor. Across three samples, we found that people who reported doing more (versus less or equal amounts) were *less* satisfied with their relationships in the moment and over time, and less satisfied with the division of labor itself. However, bearing more of the household burden was not associated with lower relationship satisfaction or declines in satisfaction over time when people felt more appreciated by their partners.

This work extends a small but growing body of research showing that feeling appreciated can offset problematic relationship dynamics that undermine satisfaction (Barton et al., 2015; Park et al., 2019) by focusing on division of labor, an issue that every cohabiting couple faces. This work also extends division of labor research. First, while research on household labor often focuses on equality versus inequality, our findings build on prior work suggesting that doing more may be more problematic than doing less (e.g., Van Willigen & Drentea, 2001). Second, aligning with recent work (e.g., Waddell et al., 2021), women contributed more, and doing more undermined relationship quality (although we did not measure traditional gender roles, which may moderate these effects). Yet critically, the beneficial effects of feeling appreciated were consistent for both women and men. Third, we extend past work showing that people feel more appreciated when contributions are more equal and fair (Klumb et al., 2006; Mikula et al., 2009), by demonstrating that feeling appreciated buffers against the relationship costs of

doing more even when those contributions are viewed as unequal and unfair. Fourth, we found evidence that feeling appreciated might also buffer people against being less satisfied when their partners do more. Future research should examine whether people who contribute less, but feel appreciated by their partners, are less prone to guilt or concerns about how their partners perceive them, or perhaps they contribute in other ways (e.g., managing familial relationships).

Is feeling satisfied despite an unequal and unfair division of labor necessarily a good thing? Feeling appreciated might help balance the egoistic tendency for people to over-perceive their contributions (Press & Townsley, 1998; Reis et al., 2018), and also help couples ride out short-term inequalities (e.g., unusually heavy work demands). However, the buffering effect of feeling appreciated might also disincentivize people from pushing for more equality when unfair division of labor is severe and chronic. While gratitude does not promote complacency (Armenta et al., 2020), a certain level of dissatisfaction may be needed to propel change. Thus, when studying prosocial processes such as appreciation, it is important to consider both their benefits and potential costs.

Are the buffering effects of feeling appreciated simply due to more satisfied people feeling more appreciated? Our correlational data cannot answer questions of causality. However, our analyses examining changes in relationship satisfaction controlled for initial levels of satisfaction. By doing so, we ruled out the possibility that these effects simply reflect pre-existing differences between more and less satisfied couples. Instead, our results appear to demonstrate that feeling appreciated plays a unique buffering role beyond couples' initial levels of satisfaction.

Constraints on Generality

Women and those who were not employed reported doing more of the household labor, but our key findings did not differ by gender, employment status, income, age, or relationship length. Nonetheless, there are potential constraints on the generality of our findings. Our samples were from two Western countries (U.S. and Canada), limiting the generalizability of our findings to other countries that may not have the same gender and relationship expectations. Second, our samples primarily consisted of two-person, mixed-gender couples, hence we caution interpreting findings across couples with different gender and relationship configurations (e.g., same-gender, polyamorous). While some research suggests same-gender couples divide labor more equitably (Shockley & Shen, 2016) and polyamorous couples may divide tasks in different ways (Schippers, 2019), we expect the benefits of feeling appreciated should generalize across different relationship types. Third, we used a self-report measure of division of labor. Time-use surveys might reveal differences in how much people are actually contributing. However, we found high rates of agreement between partners. Moreover, perceptions often shape evaluations of relationships more strongly than reality (Pollmann & Finkenauer, 2009). Hence, people's *perceptions* of their contributions, rather than actual recordings, should more powerfully influence how satisfied people feel in their relationships. Finally, our data were gathered during the COVID-19 pandemic, raising the question of generalizability across time. We did see stability in how household labor was divided across the weeks and months that we tracked our longitudinal samples (Table S3; see also, Shafer et al., 2020). Although replication outside the pandemic is needed, we have no reason to expect that feeling appreciated would *not* be a buffer for people who contribute more in a different global context.

Conclusion

In sum, results from more than 2,000 people living with a romantic partner during the COVID-19 pandemic provide robust evidence that doing more of the household labor is associated with lower relationship satisfaction. While an equal and fair distribution of labor between partners is ideal, inequality is common: more than half of our participants reported their division of labor was uneven or unfair. Our work reveals that feeling appreciated by one's partner may help people maintain relationship satisfaction over time despite feeling they contribute more. Although we might imagine someone running themselves ragged doing all the household labor and growing more resentful and discontent by the moment, these results suggest that this may only be the case when people feel taken for granted. Instead, for people with more appreciative partners, doing more may provide them with more opportunities to feel valued and recognized.

References

Ackerman, R. A., & Kenny, D. A. (2016, December). APIMPowerR: An interactive tool for Actor-Partner Interdependence Model power analysis [Computer software]. Available from https://robert-a-ackerman.shinyapps.io/APIMPowerRdis/

Adams, J. S. (1965). Inequity in social exchange. In *Advances in experimental social psychology* (Vol. 2, pp. 267–299). Elsevier.

Algoe, S. B. (2012). Find, remind, and bind: The functions of gratitude in everyday relationships. *Social and Personality Psychology Compass*, *6*(6), 455–469.

Amaro, L. M. (2017). Dyadic effects of gratitude on burden, conflict, and contribution in the family caregiver and sibling relationship. *Journal of Applied Communication Research*, *45*(1), 61–78. https://doi.org/10.1080/00909882.2016.1248464

Armenta, C. N., Fritz, M. M., Walsh, L. C., & Lyubomirsky, S. (2020). Satisfied yet striving: Gratitude fosters life satisfaction and improvement motivation in youth. *Emotion*. https://doi.org/10.1037/emo0000896

Baker, L. R., McNulty, J. K., & VanderDrift, L. E. (2017). Expectations for Future Relationship Satisfaction: Unique Sources and Critical Implications for Commitment. *Journal of Experimental Psychology. General*, *146*(5), 700–721. https://doi.org/10.1037/xge0000299

Barton, A. W., Futris, T. G., & Nielsen, R. B. (2015). Linking financial distress to marital quality: The intermediary roles of demand/withdraw and spousal gratitude expressions: Demand/withdraw and spousal gratitude expressions. *Personal Relationships*, *22*(3), 536–549. https://doi.org/10.1111/pere.12094

Berger, A. R., & Janoff-Bulman, R. (2006). Costs and satisfaction in close relationships: The role of loss-gain framing. *Personal Relationships*, *13*(1), 53–68.

Blair, S. L., & Johnson, M. P. (1992). Wives' Perceptions of the Fairness of the Division of Household Labor: The Intersection of Housework and Ideology. *Journal of Marriage and Family*, *54*(3), 570–581. https://doi.org/10.2307/353243

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213.

Craig, L., & Churchill, B. (2021). Dual-earner parent couples' work and care during COVID-19. *Gender, Work & Organization*, *28*(S1), 66–79. https://doi.org/10.1111/gwao.12497

Gordon, A. M., Impett, E. A., Kogan, A., Oveis, C., & Keltner, D. (2012). To have and to hold: Gratitude promotes relationship maintenance in intimate bonds. *Journal of Personality and Social Psychology*, *103*(2), 257–274. https://doi.org/10.1037/a0028723

Hawkins, A. J., Marshall, C. M., & Meiners, K. M. (1995). Exploring wives' sense of fairness about family work: An initial test of the distributive justice framework. *Journal of Family Issues*, *16*(6), 693.

Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., & Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research*, *18*(7), 873–880.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social Relationships and Mortality

Risk: A Meta-analytic Review. *PLoS Medicine*, *7*(7).

https://doi.org/10.1371/journal.pmed.1000316

Jackson, G. L., Trail, T. E., Kennedy, D. P., Williamson, H. C., Bradbury, T. N., &

Karney, B. R. (2016). The Salience and Severity of Relationship Problems among

Low-Income Couples. *Journal of Family Psychology : JFP : Journal of the*

Division of Family Psychology of the American Psychological Association

(Division 43), *30*(1), 2–11. https://doi.org/10.1037/fam0000158

Kansky, J. (2018). What's love got to do with it?: Romantic relationships and well-being.

Handbook of Well-Being. Salt Lake City, UT: DEF Publishers.

Klumb, P., Hoppmann, C., & Staats, M. (2006). Division of Labor in German Dual-

Earner Families: Testing Equity Theoretical Hypotheses. *Journal of Marriage and*

Family, *68*(4), 870–882. https://doi.org/10.1111/j.1741-3737.2006.00301.x

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening

scale for anxiety and depression: The PHQ–4. *Psychosomatics*, *50*(6), 613–621.

Lavee, Y., & Katz, R. (2002). Divison of labor, perceived fairness, and marital quality:

The effect of gender ideology. *Journal of Marriage and Family*, *64*(1), 27–39.

Mikula, G., Schoebi, D., Jagoditsch, S., & Macher, S. (2009). Roots and correlates of

perceived injustice in the division of family work. *Personal Relationships*, *16*(4),

553–574. https://doi.org/10.1111/j.1475-6811.2009.01239.x

Park, Y., Impett, E. A., MacDonald, G., & Lemay Jr, E. P. (2019). Saying "thank you":

Partners' expressions of gratitude protect relationship satisfaction and

commitment from the harmful effects of attachment insecurity. *Journal of Personality and Social Psychology*, *117*(4), 773.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package 'nlme.' *Linear and Nonlinear Mixed Effects Models, Version*, *3*(1).

Pollmann, M. M. H., & Finkenauer, C. (2009). Investigating the Role of Two Types of Understanding in Relationship Well-Being: Understanding Is More Important Than Knowledge. *Personality and Social Psychology Bulletin*, *35*(11), 1512–1527. https://doi.org/10.1177/0146167209342754

Press, J. E., & Townsley, E. (1998). Wives' and husbands' housework reporting: Gender, class, and social desirability. *Gender & Society*, *12*(2), 188–218.

Qian, Y., & Sayer, L. C. (2016). Division of labor, gender ideology, and marital satisfaction in East Asia. *Journal of Marriage and Family*, *78*(2), 383–400.

Reis, H. T., Lee, K. Y., O'Keefe, S. D., & Clark, M. S. (2018). Perceived partner responsiveness promotes intellectual humility. *Journal of Experimental Social Psychology*, *79*, 21–33. https://doi.org/10.1016/j.jesp.2018.05.006

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The Investment Model Scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships*, *5*(4), 357–387. https://doi.org/10.1111/j.1475-6811.1998.tb00177.x

Schippers, M. (2019). Polyamory. In *The Blackwell Encyclopedia of Sociology* (pp. 1–2). American Cancer Society. https://doi.org/10.1002/9781405165518.wbeosp047.pub2

Shafer, K., Scheibling, C., & Milkie, M. A. (2020). The Division of Domestic Labor
before and during the COVID-19 Pandemic in Canada: Stagnation versus Shifts
in Fathers' Contributions. *Canadian Review of Sociology/Revue Canadienne de
Sociologie*, *57*(4), 523–549. https://doi.org/10.1111/cars.12315

Shockley, K. M., Clark, M. A., Dodd, H., & King, E. B. (2020). Work-family strategies
during COVID-19: Examining gender dynamics among dual-earner couples with
young children. *Journal of Applied Psychology*, *106*(1), 15.
https://doi.org/10.1037/apl0000857

Shockley, K. M., & Shen, W. (2016, July 13). *Couple Dynamics*. The Oxford Handbook
of Work and Family. https://doi.org/10.1093/oxfordhb/9780199337538.013.10

ter Kuile, H., Kluwer, E., Finkenauer, C., & Lippe, T. (2017). Predicting adaptation to
parenthood: The role of responsiveness, gratitude, and trust: Predicting perceived
adaptation to parenthood. *Personal Relationships*, *24*.
https://doi.org/10.1111/pere.12202

Thielemans, G., Fallesen, P., & Mortelmans, D. (2020). Division of Household Labor and
Relationship Dissolution in Denmark 2001–2009. *Journal of Family Issues*,
0192513X20949890. https://doi.org/10.1177/0192513X20949890

van Eldik, W. M., de Haan, A. D., Parry, L. Q., Davies, P. T., Luijk, M. P. C. M., Arends,
L. R., & Prinzie, P. (2020). The interparental relationship: Meta-analytic
associations with children's maladjustment and responses to interparental conflict.
Psychological Bulletin, *146*(7), 553–594. https://doi.org/10.1037/bul0000233

Van Willigen, M., & Drentea, P. (2001). Benefits of equitable relationships: The impact of sense of fairness, household division of labor, and decision making power on perceived social support. *Sex Roles*, *44*(9), 571–597.

Van Yperen, N. W., & Buunk, B. P. (1990). A longitudinal study of equity and satisfaction in intimate relationships. *European Journal of Social Psychology*, *20*(4), 287–309. https://doi.org/10.1002/ejsp.2420200403

Waddell, N., Overall, N. C., Chang, V. T., & Hammond, M. D. (2021). Gendered division of labor during a nationwide COVID-19 lockdown: Implications for relationship problems and satisfaction. *Journal of Social and Personal Relationships*, 0265407521996476. https://doi.org/10.1177/0265407521996476

Table 1. Simple slopes for linear and curvilinear effects of perceived household DoL on concurrent relationship satisfaction, expected future relationship satisfaction, and actual changes in relationship satisfaction at low (-1SD) and high (+1SD) levels of feeling appreciated.

DV: Relationship Satisfaction	IV	Low Appreciated (-1SD)						High Appreciated (+1SD)					
		B	*t*	*p*	*LL*	*UL*	*r*	*B*	*t*	*p*	*LL*	*UL*	*r*
Predicting Concurrent Relationship Satisfaction at Baseline													
Sample A - Baseline	DoL	0.00	-0.02	0.98	-0.10	0.10	0.00	-0.02	-0.49	0.62	-0.10	0.06	0.02
	DoL²	-0.12	-3.42	0.001	-0.19	-0.05	0.10	0.02	0.53	0.60	-0.05	0.08	0.02
Sample B - Baseline	DoL	0.02	0.32	0.75	-0.10	0.13	0.01	0.00	-0.04	0.97	-0.11	0.11	0.00
	DoL²	-0.08	-1.96	0.05	-0.16	0.00	0.08	-0.01	-0.12	0.90	-0.09	0.08	0.01
Sample C - Baseline	DoL	0.05	0.70	0.49	-0.09	0.18	0.04	-0.11	-1.41	0.16	-0.26	0.04	0.09
	DoL²	-0.09	-1.41	0.16	-0.22	0.04	0.07	0.06	0.70	0.48	-0.11	0.22	0.04
Predicting Concurrent Relationship Satisfaction in Weekly and Follow-ups Assessments													
Sample B - Follow-ups (T2-T5)	DoL	0.07	2.21	0.03	0.01	0.14	0.14	-0.05	-1.57	0.12	-0.12	0.01	0.09
	DoL²	-0.06	-2.30	0.02	-0.11	-0.01	0.14	0.01	0.49	0.63	-0.04	0.07	0.02
Sample C - Weekly (W1-W3)	DoL	0.11	1.53	0.13	-0.03	0.26	0.15	-0.06	-0.77	0.44	-0.21	0.09	0.06
	DoL²	-0.10	-1.79	0.08	-0.21	0.01	0.14	0.02	0.38	0.71	-0.10	0.15	0.03
Sample C - Follow-up (6 Month)	DoL	0.04	0.35	0.73	-0.17	0.24	0.02	-0.13	-1.38	0.17	-0.32	0.06	0.09
	DoL²	-0.27	-2.82	0.01	-0.45	-0.08	0.17	0.00	0.01	1.00	-0.17	0.17	0.00
Predicting Expected Relationship Satisfaction													
Sample A - Baseline	DoL	-0.12	-1.01	0.31	-0.34	0.11	0.03	-0.05	-0.62	0.54	-0.23	0.12	0.02
	DoL²	-0.33	-4.10	<0.001	-0.49	-0.17	0.12	-0.08	-1.06	0.29	-0.24	0.07	0.03
Sample B - Baseline (T1)	DoL	0.05	0.40	0.69	-0.21	0.32	0.02	-0.17	-1.35	0.18	-0.42	0.08	0.06
	DoL²	-0.12	-1.24	0.21	-0.32	0.07	0.05	-0.07	-0.70	0.49	-0.28	0.13	0.03
Sample B - Follow-ups (T2-T5)	DoL	0.12	1.70	0.09	-0.02	0.27	0.08	0.09	1.29	0.20	-0.05	0.23	0.06
	DoL²	-0.01	-0.17	0.87	-0.13	0.11	0.01	0.01	0.09	0.93	-0.12	0.13	0.01
Predicting Change in Relationship Satisfaction Across Time													
Sample B - Follow-ups (T2-T5)	DoL	0.13	3.53	<0.001	0.06	0.20	0.18	-0.02	-0.58	0.56	-0.09	0.05	0.03
	DoL²	-0.02	-0.53	0.59	-0.08	0.05	0.04	-0.01	-0.32	0.75	-0.08	0.06	0.01
Sample C - Weekly (W1-W3)	DoL	0.07	1.21	0.23	-0.04	0.17	0.06	0.01	0.24	0.81	-0.10	0.13	0.01
	DoL²	-0.12	-2.92	0.01	-0.21	-0.04	0.10	0.02	0.40	0.69	-0.08	0.12	0.01
Sample C - Follow-up (6 Month)	DoL	-0.03	-0.32	0.75	-0.21	0.15	0.02	-0.09	-1.03	0.30	-0.25	0.08	0.07
	DoL²	-0.28	-3.37	0.01	-0.44	-0.12	0.20	0.01	0.18	0.86	-0.13	0.16	0.01

Note: Linear and curvilinear main effects of DoL included in all models. For main effects and interactions, see Tables S4-S7. $r = \sqrt{t^2/(t^2+df)}$
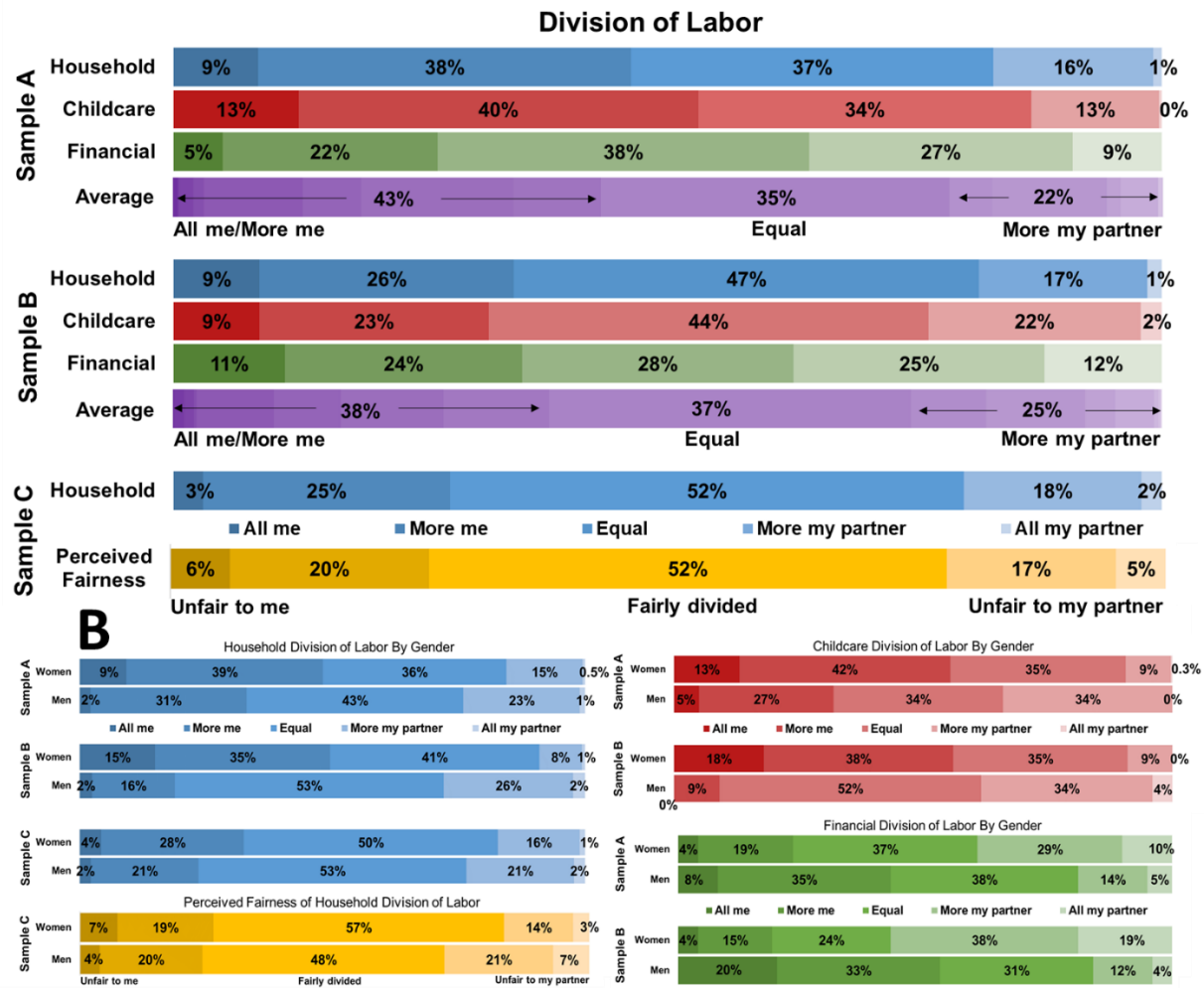
Figure 1. Breakdown of perceived division of labor (DoL) at baseline in all three samples. A: perceived DoL and perceived fairness of household DoL since pandemic began, B: DoL by gender. Average DoL aggregates household, childcare (when relevant), and financial contributions (includes fractional values). Purple bars are shaded based on actual values; percentages represent "All me/More me" (-2 to -.33), "Equal" (0) and "More my partner" (0.33 to 1.67; no participants had average scores equal to "All my partner").
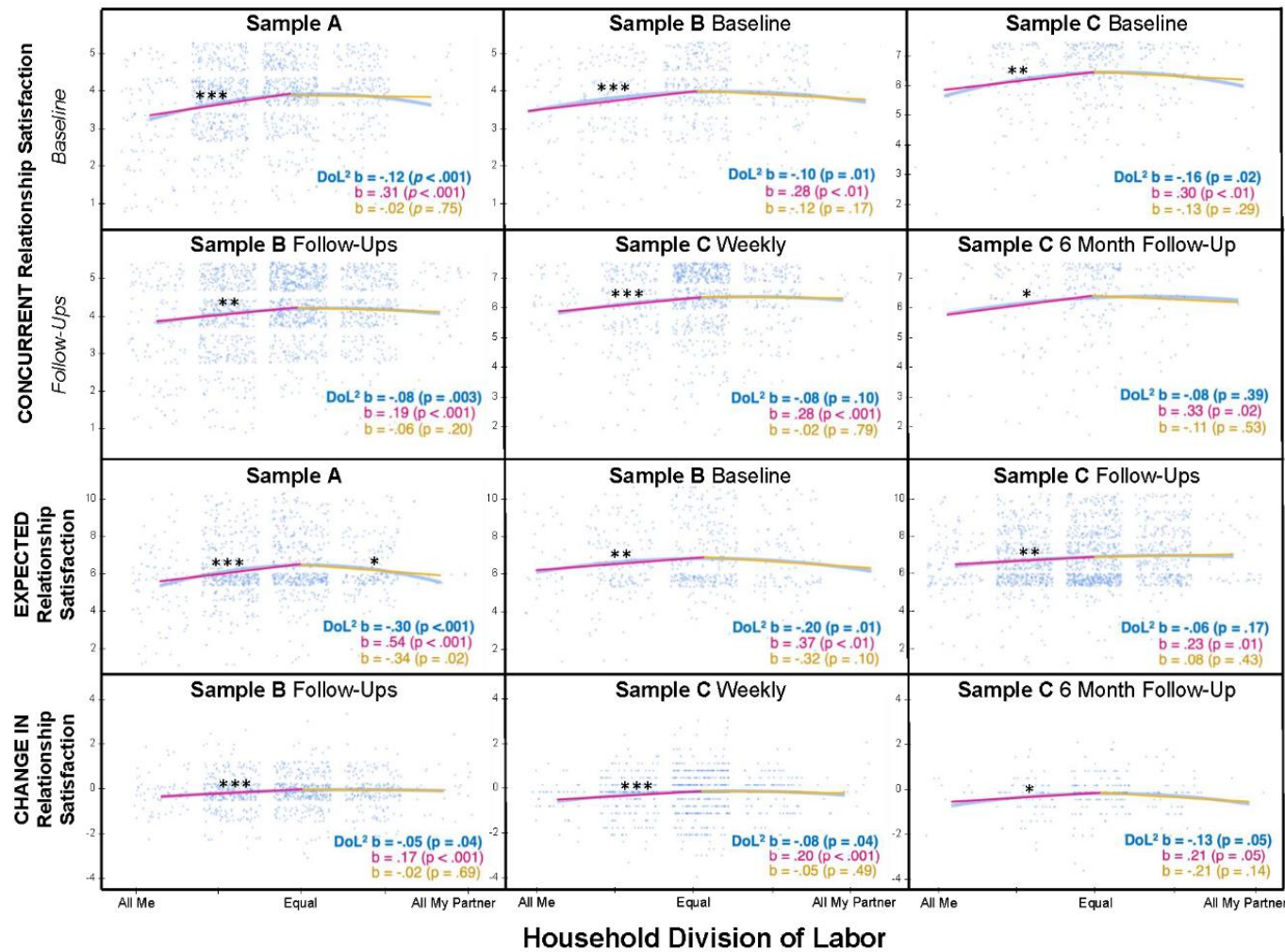
Figure 2. Curvilinear associations between household DoL and relationship satisfaction across samples and measures of relationship satisfaction. Piecewise regressions illustrate significant linear slopes from "all me" to "equal" and generally non-significant linear slopes from "equal" to "all my partner." B's are unstandardized estimates. *** = $p \leq .001$, ** $\leq .01$, * $\leq .05$, $^f \leq .10$. Sample B follow-ups = four follow-ups every 3 months. Sample C weekly = three weekly follow-ups.
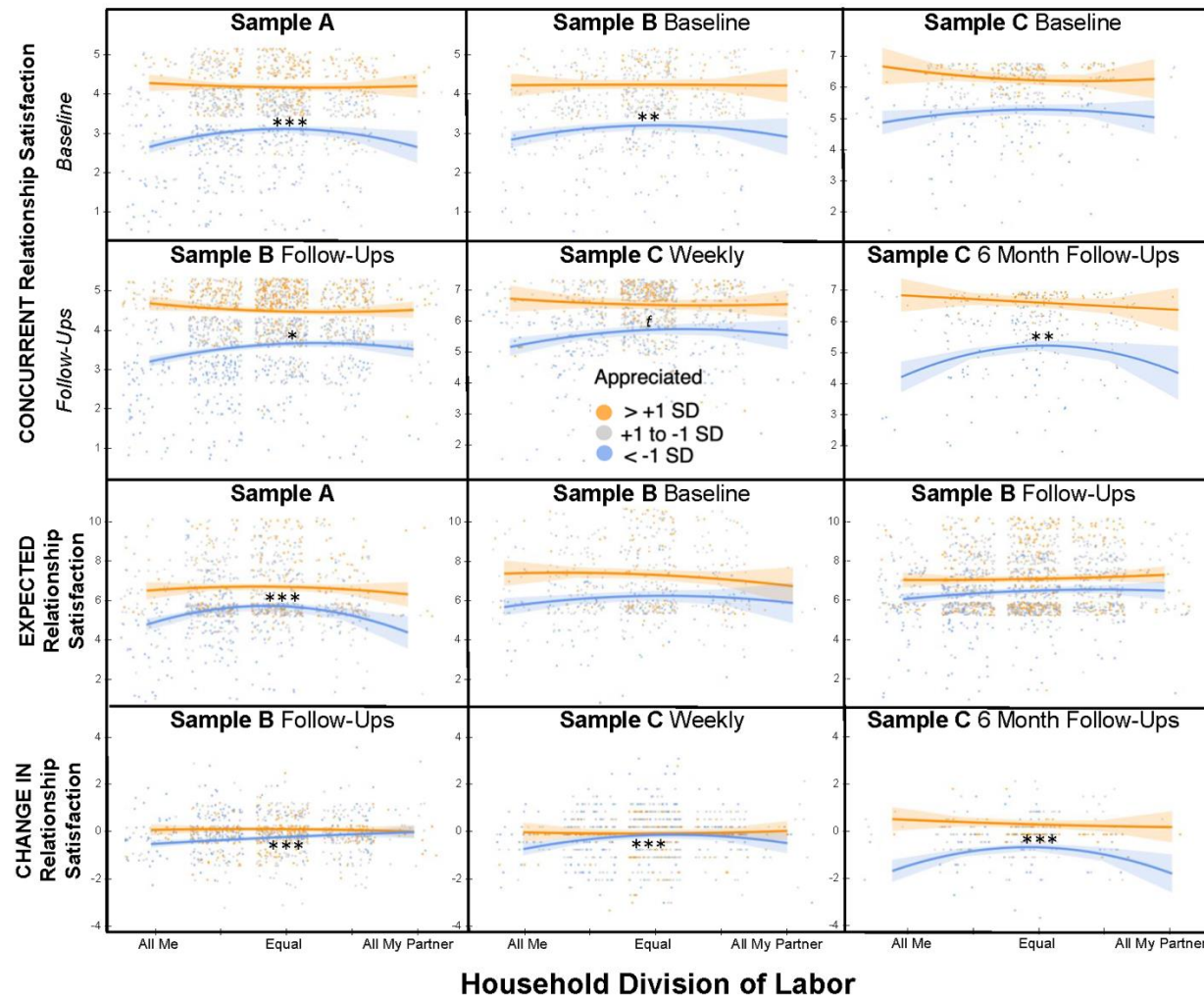
Figure 3. Feeling appreciated moderating the association between perceived household DoL and relationship satisfaction across samples and measures of relationship satisfaction. *** = $p \leq$ .001, ** $\leq$ .01, * $\leq$ .05, $^f \leq$ .10, for linear and/or curvilinear simple slopes at -1 *SD* appreciated. Slopes were nonsignificant at +1 *SD* appreciated. Tables S6-9 includes estimates for main effects and moderations. Table 1 includes simple slopes for curvilinear moderations.