Cohort Trends in the Gender Distribution of Household Tasks in the United States and the Implications for Understanding Disability

Connor Sheehan, PhD[1], **Benjamin W. Domingue, PhD**[2], **Eileen Crimmins, PhD**[3]

[1]Arizona State University, Tempe, USA

[2]Stanford University, CA, USA

[3]University of Southern California, Los Angeles, USA

Abstract

Objectives: Measures of disability depend on health and social roles in a given environment. Yet, social roles can change over time as they have by gender. We document how engagement in Instrumental Activities of Daily Living (IADLs) is shifting by gender and birth cohort among older adults, and the challenges these shifts can create for population-level estimates of disability.

Method: We used the Health and Retirement Study ($N = 25,047$) and multinomial logistic regression models with an interaction term between gender and birth cohort to predict limitation and nonperformance relative to no difficulty conducting IADLs.

Results: Nonperformance of IADLs have significantly decreased among younger cohorts. Women in younger cohorts were more likely to use a map, whereas men in younger cohorts were more likely to prepare meals and shop.

Discussion: Failing to account for gender and cohort changes in IADL, performance may lead to systematic bias in estimates of population-level disability.

Keywords

disability; cohort; gender; household labor; instrumental activities of daily living

Introduction

The aging of the U.S. population (Vincent & Velkoff, 2010) will have many consequences, including the restructuring of public health priorities (Auerbach et al., 2017). One priority likely to receive more attention is population-level predictors of healthy aging and prolonged independent living (Crimmins, Hayward, Hagedorn, Saito, & Brouard, 2009; Shaw et al.,

2017; Thomeer, Mudrazija, & Angel, 2016). Identifying predictors of healthy aging and independent living depends on accurately measuring health and disability, both challenging tasks (Sheehan & Tucker-Drob, 2017; Verbrugge & Jette, 1994). Disability is particularly difficult to measure. This difficulty stems from its definition as the inability to perform social roles or expectations in a given environment (Verbrugge & Jette, 1994). That is, measuring disability depends not only on underlying health capacity but also on the social roles, expectations, and the broader environment (Crimmins et al., 2009; Freedman, Martin, & Schoeni, 2002; Schoeni, Martin, Andreski, & Freedman, 2005; Verbrugge & Jette, 1994). Social roles (McRobbie, 2009; Ryder, 1965) and environments (Steinfeld & Danford, 1999) are constantly shifting, this implies that disability is assessed and reassessed within new social and environmental circumstances.

One commonly employed measure of disability is the Instrumental Activities of Daily Living (IADLs) battery (Crimmins & Saito, 2000; Hybels, Pieper, Blazer, Fillenbaum, & Steffens, 2010; Millán-Calenti et al., 2010; Palmer, Espino, Dergance, Becho, & Markides, 2012; White-Means & Rubin, 2008; Yong, Saito, & Chan, 2010). This battery is common in gerontological research as its questions are designed to assess an individual's ability to perform activities required for independent living (Spector, Katz, Murphy, & Fulton, 1987). Specifically, the questions of this battery gauge the ability to independently conduct basic household activities such as using a map, using a telephone, managing money, taking medication, shopping for groceries, and preparing meals. Consistent with other health research, previous research using this battery has generally shown that women report greater levels than men do of health limitations in IADLs (Crimmins et al., 2009; Crimmins & Saito, 2000; Crimmins, Kim, & Solé-Auró, 2011; Fleishman, Spector, & Altman, 2002; Murtagh & Hubert, 2004; Sheehan & Tucker-Drob, 2017). These findings, however, are complicated by the notion that performance and engagement in IADLs are influenced not only by health (Geist & Tabler, 2018) but also by differing expectations for men and women in performing these activities and household labor more broadly (Fleishman et al., 2002; Sheehan & Tucker-Drob, 2017).

Previous research has implicitly assumed that differences between men and women in engaging and performing IADLs are static and have not changed in recent decades. Yet, the distribution of these tasks between men and women has become more egalitarian compared with previous time periods (e.g., Becker, 1985). The gradual shift toward a more egalitarian distribution of domestic labor has been documented as a social process that has unfolded largely by birth cohort (Artis & Pavalko, 2003; Brooks & Bolzendahl, 2004; Fuwa, 2004; McRobbie, 2009; Percheski, 2008; Teerawichitchainan, Knodel, Loi, & Huy, 2010). There are broad social and economic processes driving this cohort-based shift (Calasanti & Bailey, 1991; Carlson & Lynch, 2017). For example, female employment has increased steadily by birth cohort (Percheski, 2008), and marriages in which the female is in the labor force tend to have a more egalitarian split of domestic labor (Nitsche & Grunow, 2016). Although there is substantial research documenting demographic differences in health limitations for IADLs by gender and cohort (Murtagh & Hubert, 2004; Sheehan & Tucker-Drob, 2017; Spillman, 2004; Taylor & Lynch, 2011; Verbrugge, Brown, & Zajacova, 2017), such research has not accounted for cohort changes in the expectation and fulfillment of household tasks by men and women and how such changes may be reflected in reports of IADLs and IADL

limitations. Our analysis documents how links between gender and IADL fulfillment are not static but have changed significantly—a process driven by cohorts.

Given that expectations regarding household labor for men and women have changed over time (Artis & Pavalko, 2003) and that differential reporting of task performance due to these changing expectations has biased estimates of IADL differentials (Sheehan & Tucker-Drob, 2017), this is an important issue that has implications for measuring population-level disability (Freedman et al., 2002, Freedman et al., 2013). The potential bias arises from the fact that many IADL batteries, such as that used in the Health and Retirement Study (HRS), allow respondents to report not doing an IADL for reasons other than health or memory impairment. If these "do not do" responses vary by gender or cohort, then these responses may *systematically* bias substantive findings regarding health-related IADL limitations. We define these "do not do" responses as nonperformance, as these respondents are not performing the activity for reasons other than health. Different interpretations of these responses can result in diametrically opposed findings regarding IADL prevalence and severity over time and by gender.

Whether gender and cohort changes in IADL task performance are affecting bias in estimates of IADL functioning problems is unknown. Given that household domestic labor is becoming more egalitarian by cohort, we anticipate that gendered reports of nonperformance of IADLs have decreased among younger cohorts. We seek to document how responses to the IADL questions, particularly nonperformance responses, have changed by gender and birth cohort. We also illustrate how reports of nonperformance vary by other demographic characteristics.

Method

Data

We used data from the HRS, a nationally representative panel study of Americans at least 51 years old and their spouses, conducted biennially since the early 1990s. Extensive previous research has used the HRS to document trends and differences in IADLs (e.g., Chen & Sloan, 2015; Freedman, Aykan, Wolf, & Marcotte, 2004; Freedman et al., 2013; Henning-Smith, Shippee, & Capistrant, 2017; Nguyen, Rist, & Glymour, 2016). Specifically, we use the Health and Retirement Study, ([RAND HRS "O File"]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (Grant Number NIA U01AG009740). Ann Arbor, MI.

We used six cohorts of the HRS comprising more than 35 birth years, and coded gender male (=1) and female as the reference (=0). As labeled by the HRS (sometimes for the survey in which data on them were first gathered), the six cohorts of data we used are Asset and Health Dynamics Among Oldest Old (AHEAD; born before 1924, coded "0"), Children of the Depression (CODA; born 1923–1930, coded "1"), HRS (born 1931–1941, coded "2"), War Babies (WB, born 1942–1947, coded "3"), Early Baby Boomers (EBB, born 1948–1953, coded "4"), and Mid-Baby Boomers (MBB, Born 1954–1959, coded "5"). For the HRS and AHEAD cohorts, we used data from the 1995/1996 wave as this was when the HRS first implemented the harmonized IADL questionnaire it has used since. For other

cohorts, we used observations from their initial HRS wave (Table 2 indicates the first wave for each cohort we analyzed). Focusing on the first observation of IADLs for each cohort minimized problems associated with mortality selection (Domingue et al., 2017; Zajacova & Burgard, 2013) and made the age profiles between cohorts as similar as possible. We also calculated IADL prevalence data, for the entire older adult (51+) population (and thus the entire HRS samples for those years), using the 2000 and 2010 HRS waves to demonstrate that different strategies for handling the nonperformance responses can lead to different substantive results.

We tested for cohort differences using multivariable models, we compared responses to IADL items for individuals at their first wave in the HRS (aside for the HRS and AHEAD cohorts). We dropped the small number of respondents with any missing IADL questions, respondents who were spouses born after 1959, and those in the HRS and AHEAD waves who were lost to follow up before 1995/1996. This yielded an analytical sample of 25,047. An argument could be made to examine only the married heterosexual couples, but we examined the entire sample to be consistent with previous research that aimed at examining population-wide trends in IADLs (e.g., Freedman et al., 2013). Although we found that limiting our sample to only those who were married strengthened our results, our overall goal was to evaluate the influence of nonperformance for the entire population.

As mentioned, because we did not use the HRS and AHEAD responses before the 1995/1996 wave and because the CODA wave entered the study at an older age, their initial responses were given at older ages than those for the WB, EBB, and MBB waves. Nevertheless, we obtained similar results to those we will show when focusing on the three cohorts with similar age (see Table 2 which shows cohort declines in nonperformance among the youngest cohorts and Supplemental Table 2 which shows larger differences in the intercepts than slopes). Finally, for multivariable models below, we imputed the small number of missing values for demographic characteristics (of which marital status, with 0.002% missing, had the highest missing rate) using Stata's multiple imputation suite. Results (not shown) from models that used listwise deletion were consistent with the imputed models.

Measures

We coded the HRS IADL questions to be consistent with previous research (Sheehan & Tucker-Drob, 2017). The HRS asked respondents a series of questions about whether due to a health or memory problem they had difficulty using a map, using a telephone, managing money, taking medicine, shopping for groceries, or preparing meals. We coded the responses as (a) no difficulty conducting the activity (base outcome), (b) having difficulty or inability in performing the activity due to health or memory problems, or (c) not performing the activity for reasons other than health or memory. In addition, and consistent with previous research (Sheehan & Tucker-Drob, 2017), we coded those who report not taking medication (4.1%) as having no difficulty.

As mentioned, we coded our cohorts in a linear specification based on their HRS birth year designations. We also tested a specification of cohorts based on 5-year birth year intervals and another based on continuously specified birth year; these yielded similar results (not

shown) to our analyses based on HRS designated cohorts. In ancillary analyses, we also found that the linear specification of cohorts fit better than other specifications (such as semiparametric). The linear specification was also consistent with the notion of increasing gender egalitarianism by cohort.

In our multivariate models, we adjusted for demographic factors that vary by cohort. We controlled for age measured in years, centered at age 51 (i.e., respondent age—51). We coded educational attainment as less than high school (referent), high school, and greater than high school. Controlling for education in analyses was important because not only are there substantial differences in IADL disability by educational attainment (Bound, Schoenbaum, & Waidmann, 1995), there are also important cohort-based differences in other health outcomes (Hayward & Sheehan, 2016; Masters, Hummer, & Powers, 2012). Also, it was unclear whether nonperformance of IADLs may vary by educational attainment. Controlling for race and Hispanic ethnicity was necessary because the racial and ethnic composition of cohorts has changed over time and because IADL disability varies by these characteristics (Melvin, Hummer, Elo, & Mehta, 2014). The race and ethnicity categories we used were non-Hispanic White (referent), non-Hispanic Black, foreign-born Hispanic, native-born Hispanic, and non-Hispanic Other. Finally, because nonperformance of IADLs may vary by marital status as respondents rely on partners to conduct activities, we coded marital status as married (referent), divorced/separated, widowed, and never married. We show results from the fully adjusted models, including controls in Supplemental Table 3; below, we present results for models using an interaction term for gender and HRS cohort.

Statistical Approach

We first present a descriptive analysis that documents the implications of handling nonperformance reports for the population. We calculated prevalence of IADLs based on three different ways of handling reports of nonperformance. We then document age trajectories by cohort and gender among all respondents younger than 90 (120,093 person-waves observations) from the 1995/1996 wave through the 2010 wave. The upper age cutoff was set at 90 given previous research on mortality selection in the HRS (Zajacova & Burgard, 2013). In this figure where the trajectories are presented, we considered the proportion of nonperformance and limited observations as a function of cohort and gender.

To formally test the cohort trend, we fit multinomial logistic regression models that predicted the responses for each IADL question. Specifically, we fit multinomial models that predicted whether the respondent reported being able to conduct the IADL without limitation (base outcome) relative to being limited or unable to conduct it and relative to not performing it. We used interaction terms between gender and cohort to test gender differences in cohort trends in health and egalitarianism. These models controlled for age at survey, race and Hispanic origin, educational attainment, and marital status. For ease of interpretability, we present results from the multinomial models in exponentiated form as Relative Risk Ratios (RRR).

Sensitivity Analyses

We considered several sensitivity analyses. Age and period effects are alternative mechanisms that may complicate our suggestion that interpretation of IADLs need be cohort based. It is possible that as individuals age, they become more inclined to report not doing each of the activities or that cohorts shift their responses over time. To explore this, we graphically documented the age trajectories by cohort in Figure 1. As mentioned, we found small shifts in cohort reports of nonperformance as they aged, aside from the oldest cohort using a map. To evaluate potential period effects, we document the proportion of the respondents in each wave who reported nonperformance from the activity; these figures are shown as Supplemental Figures 5 to 7. In general, reports of nonperformance were constant across the HRS waves within cohort. Age or period effects did not seem responsible for the declines in reports of nonperformance in the IADL items. We also fit models with interactions between all the controls and gender and cohort which provided substantively similar results, indicating the gender cohort interaction was significant accounting for interactions between gender, cohort, and the other covariates.

Results

Treatment of Nonperformance IADL Responses Alters National Prevalence and Severity Results

Table 1 shows estimates of the prevalence of IADL disability under three strategies for handling nonperformance responses: (a) treating such responses as missing, (b) treating them as not indicating a limitation in the activity, and (c) treating them as indicating a limitation in the activity. Under the first strategy, excluding nonperformance responses, IADL prevalence and severity increased for both men (4.7% increase) and women (3.9% increase) between 2000 and 2010. Under the second strategy of treating nonperformance response as not indicating a limitation, the prevalence of IADL limitations increased for women (1.2% increase), albeit at a smaller rate, and decreased for men (2.2% decrease). Under the third strategy of treating nonperformance as indicating a limitation, IADL prevalence and severity decreased for both men (8.9% decrease) and women (4.7% decrease) but did so more for men. Put another way, not only does the strategy for treating IADL nonperformance affect levels of prevalence and severity of limitations, it also affects findings on whether gender differences in IADL limitation have decreased or remained steady. More detailed information regarding Table 1, including nonperformance for each item by wave, is available in the Online Supplement. These results suggest that reports of nonperformance, and how they are changing, are important to understand.

Patterns of Engagement Are Largely Structured by Birth Cohort, Whereas Limitation Patterns Are Largely Structured by Age

Figure 1 depicts reports of nonperformance (left column) and health-based limitations (right column) of an IADL as a function of age, cohort, and gender. Points represent cohort-specific means by gender for those at a given age. Based on data from each birth cohort, we show LOESS curves with 95% confidence intervals for reports of nonperformance or limitation as a function of age. These results suggest that, as expected, health-based

limitations of IADLs increase with age. That is, as the older cohorts age, the proportion of its respondents reporting a limitation increases.

Conversely, for nonperformance, there were more pronounced cohort differences. This was especially the case for women using a map and for both shopping for groceries and preparing meals for men. For women using a map, there was little age difference in nonperformance among cohorts (with the exception of the AHEAD, or oldest, cohort). Instead, women from younger cohorts reported lower levels of nonperformance using a map than women in older cohorts. A similar pattern emerged for men shopping and preparing meals. Men in older cohorts reported higher levels of nonperformance than did those in younger cohorts, with little within cohort change as the cohorts aged. Overall, these results indicated cohort differences in nonperformance stable levels of nonperformance as cohorts age.

At First Observation in the HRS, There Are Substantial Differences in Nonperformance Across Birth Cohorts

Table 2 depicts the distribution of the responses to each individual IADL question stratified by gender and cohort upon first entrance in the HRS. There were substantial differences in nonperformance of individual IADLs. For women, the proportion who reported nonperformance in using a map is 30.7% for the AHEAD cohort and 17.0% for the CODA cohort but less than 10% for the three most-recent cohorts. Similarly, the proportion of women who reported nonperformance in managing money was 3.3% for the AHEAD cohort but less than 1.5% for each of the three most-recent cohorts, and below 1.0% at entry for two of the three most-recent cohorts. For shopping, 4.3% of women in the AHEAD cohort reported nonperformance, but less than 1% of women in all other cohorts did so.

For men, the cohort differential in nonperformance was more striking. The proportion of men in the AHEAD cohort who reported nonperformance using a map was 13.0% in the AHEAD cohort, but only 7.0% in the CODA cohort and less than 2.5% for the three most-recent cohorts. There were, however, still some differences between men and women using a map in the youngest cohort. Among MBB respondents, for example, only 1.5% of men but 5.9% of women reported nonperformance. Nonperformance in managing money was also greater for men among older cohorts than younger cohorts. Among the largest decreases in nonperformance by cohort for men were in with traditionally high levels of female engagement: shopping for groceries and preparing meals. Of men in the AHEAD cohort, 7.1% reported not grocery shopping, compared with only 0.6% of men in the MBB cohort. For preparation of meals, 17.3% of men in the AHEAD cohort and 5.7% in the WB cohort reported nonperformance, but less than 2.0% of men in the EBB and MBB cohorts did so.

These results are consistent with previous research suggesting nonperformance of these activities varies by gender (Sheehan & Tucker-Drob, 2017). However, they also suggest that although gender-based expectations for these activities may have diminished over time, these expectations have not fully disappeared. Indeed, although the absolute differences have narrowed, even in the youngest cohort, men were still three times more likely to report not shopping and four times more likely to report not preparing meals. Figure 2 graphically illustrates the proportion of men who reported nonperformance in each IADL. For preparing

meals, using a map, going shopping, and managing money, there were steep decreases of not engaging in the activities by birth cohort.

Table 3 documents RRRs from multinomial models that predicted health-based limitation or nonperformance relative to having no difficulty conducting the activity. We fit models for each of IADL item. To formally test how gender egalitarianism was shifting by cohort, we fit an interaction between gender and HRS birth cohort (full results, including controls, are in Supplemental Table 3). For ease of interpretability, we present the predicted cohort probabilities by gender in Supplemental Figures 1 to 3. For using a map, men were much less likely to report nonperformance than women (reference) were, at least for the oldest cohort (or the cohort intercept, RRR: 0.29, $p > .001$). The negative cohort slope for women was significant (RRR: 0.85, $p > .001$), but the gender differences in the cohort slope were not significant (see Supplemental Figure 1). Our models controlled for race/ethnicity, age, educational attainment, and marital status, we note that racial and ethnic minorities (relative to non-Hispanic Whites), those with low educational attainment (relative to those with high school or more education), and the unmarried (relative to the married) were also significantly more likely to report not using a map (relative to having no difficulty doing the activity) than their counterparts (results shown in Supplemental Table 3).

There were also significant differences in nonperformance among IADLs with nongendered expectations: managing money and using a telephone (full results shown in Supplemental Table 3). Those with less than a high school education (relative to those with high school or more education) and foreign-born Hispanics (relative to Non-Hispanic Whites) had significantly greater odds of reporting not managing money than having no difficulty managing money. Not surprisingly, those who were divorced, widowed, or never married were significantly less likely to report nonperformance managing their money than respondents who were married. Men had a significantly higher cohort intercept (than women) for reporting not using a phone but the cohort slope did not differ significantly between men and women. In other words, men from the oldest cohort had much higher odds of reporting not using a phone relative to women relative to having no difficulty using a phone. The full results also illustrated that racial/ethnic minorities (compared with non-Hispanic Whites) and those with less than a high school education (relative to those with high school or more education) had significantly higher odds of reporting not using a telephone relative to having no difficulty using a telephone.

Previous research (Sheehan & Tucker-Drob, 2017) has shown substantial gender differences in reports of preparing meals and shopping for groceries, with men more likely to indicate nonperformance. For preparing meals, we found that men had a significantly higher cohort intercept than women (RRR: 5.48, $p > .001$; also see Supplemental Figure 2). Indeed, the gender stratified models indicated positive and nonsignificant slopes for women (RRR: 1.26, not significant) and negative and significant slopes for men (RRR: 0.68, $p > .001$). The significant interaction term indicated the slopes differed by statistically significant amounts and that the negative nonperformance slope for men is significantly different than the slope for women. For shopping, we found no significant difference in the cohort intercept between men and women (RRR: 1.10, not significant; also see Supplemental Figure 3). Stratified models indicated that the slope was positive but nonsignificant for women (RRR: 1.21, not

significant), but the slope was negative and significant for men (RRR: 0.71, p >.001). The interaction term indicated significant gender differences in the cohort slope of nonperformance. This was due to the fact that women were less likely to report nonperformance of shopping and men became less likely to report nonperformance in younger cohorts. One important implication from the results of the models in Table 3 is that the potential bias introduced by the nonperformance responses has subsided by cohort (see Supplemental Figure 4). Overall, these results stress not only gender differences in the propensity to report not engaging in specific IADLs, but that these gender differences are becoming muted over time.

Discussion

As the American population ages, developing measures that can succinctly and accurately gauge population health, independent living, and disability is increasingly important. In this analysis, we document nonhealth-related cohort-based shifts in responses to an oft-utilized measure of disability and independent living: questions on IADL. We built on previous research by documenting how cohort shifts in engagement in household labor by men and women (Brooks & Bolzendahl, 2004) has influenced the traditionally gender-based propensity to report engaging in queried IADLs (Sheehan & Tucker-Drob, 2017). We also evaluated and found little evidence for potential age and period effects. Our results illustrate how broad social and economic changes, above and beyond any health differences (Calasanti & Bailey, 1991) can influence population-level estimates of disability among older adult Americans.

The implications of nonperformance responses are important for population estimates of disability. Different approaches to handling these responses can lead to differing estimates of IADL limitations as well as their trends and differences between men and women. This makes understanding reports of nonperformance critically important. Although health-based limitations increased with age, we found little age-related change in nonperformance in the activities. Instead, the changes in nonperformance were largely driven by birth cohort differences. One exception was the increase in reports of nonperformance in using a map, which did increase with age for men and women, likely because of the decline in driving at the oldest ages (Edwards, Lunsman, Perkins, Rebok, & Roth, 2009). Overall by cohort, women became much less likely to report nonperformance for using a map and men became less likely to report nonperformance for grocery shopping and preparing meals. Still, even in the youngest cohort men were three times more likely to report not shopping and four times more likely to report not preparing meals than women are.

Consistent with the cohort decline in reports of nonperformance, we showed that how interpretation of nonperformance can affect reports of IADL limitations has decreased considerably among more recent birth cohorts (see Table 3 and Supplemental Figure 1). The multinomial models that accounted for important covariates largely echoed the descriptive findings. Indeed, we found significant negative cohort slopes for men and women in not using a map and for men in preparing meals and shopping even after adjusting for age, race/ethnicity, educational attainment, and marital status. These results were consistent with previous research that suggests household labor is becoming more egalitarian between men

and women (Artis & Pavalko, 2003). This suggests that, assuming increasing gender egalitarianism in household tasks, going forward, there should be less potential bias in IADL measures. Still, we urge researchers to document the extent (or bounds) of bias that nonperformance reports can create, especially, when analyzing older cohorts, comparing cohorts, or making gender comparisons. Specifically, we suggest documenting the results under the assumptions that the reports of nonperformance were limited and not limited. Other alternatives are to use more performance-based measures, functional limitations, or Activities of Daily Living (ADLs) that have little gendered expectations of engagement. Indeed, in ancillary analyses, we found very few reports of nonperformance in ADLs (results not shown). Researchers should also acknowledge that these IADLs in particular may be influenced by technological advancement, indeed respondents in 2018 likely interpret "using a map" considerably differently than those in 1998 did.

We also found significant differences by race/ethnicity, educational attainment, and marital status in IADL engagement. Although documenting these differences was not our overall objective, we are unaware of other research which has done so. Specifically, we found racial and ethnic minorities had higher odds of reporting nonperformance for using a map or a telephone. Respondents with less than a high school education had significantly higher odds of reporting nonperformance in each activity except shopping. Those who were divorced, widowed, or never married were generally less likely to report nonperformance than those who were married, likely because the married could rely on a partner for many activities. These results suggest that, like that on gender or cohort differences, research on IADL differences by race, education, or marital status should also carefully gauge the sensitivity of findings and how they may be altered by nonperformance reports as these reports may systematically vary by the demographic characteristics under study.

There are some limitations to our findings. Three of the cohorts—AHEAD, CODA, and HRS—were older than the other three cohorts. Still, even in the youngest three cohorts with similar age compositions, we found declines in reports of nonperformance as well as in gender/cohort differences (see Table 2 and Supplemental Table 2). Our sensitivity analyses also showed little period decline in reports of nonperformance but large cohort differences (See Figure 1 and Supplemental Figures 4 to 6). Although we did not have a direct measure of gender egalitarianism, the increasing participation of women using maps and men preparing meals and shopping is consistent with increasing gender egalitarianism in household tasks. Our results could also be biased by mortality selection. Within the HRS, mortality selection has left the survivors advantaged relative to their full birth cohort (Zajacova & Burgard, 2013). Such selection may lead to subtle biases in analyses across cohorts (Domingue et al., 2017). We attempted to minimize this potential bias by using the earliest possible wave. In addition, previous research has suggested cohort differentials in health and increases in obesity (Hayward & Sheehan, 2016). By using multinomial models predicting health-related limitations and nonperformances, we do not feel that cohort differences in health challenge our conclusions. Finally, we did not limit our sample to partnered or heterosexual respondents which likely makes our results conservative.

Although the degree to which men and women differ in performing IADL tasks appears to be decreasing in the United States, similar declines may not have occurred elsewhere.

Indeed, previous research has documented systematic cross-national differences in the division of household labor for men and women (Davis & Greenstein, 2004). This is especially important given global population aging (Hayutin, 2007) and the design of the HRS has been used to collect similar data on aging in other countries, such as the Mexican Health and Aging Study or the Survey of Health, Aging and Retirement in Europe (Börsch-Supan & Jürges, 2005; Wong, Michaels-Obregon, & Palloni, 2017). Future research should analyze any potential bias in reports of IADLs by country, especially, in countries with varying degrees of gender egalitarianism in household labor expectations, and be cautious when making cross-national comparisons.

Disability is difficult to measure at the population level as it integrates health, technology, household structure, expectations, the environment, gender roles, and other factors. Overall, we recommend that researchers carefully conceptualize the disability measures they utilize and explicitly consider and examine how change in social expectations of engaging in queried activities can influence results. As society continues to change rapidly in its social roles, expectations, environments, technology, other characteristics, using existing measurements of disability accurately while developing new measures will become increasingly important to understanding the health of an aging population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Artis JE, & Pavalko EK (2003). Explaining the decline in women's household labor: Individual change and cohort differences. Journal of Marriage and Family, 65, 746–761.

Auerbach AJ, Charles KK, Coile CC, Gale W, Goldman D, Lee R,…Wong R (2017). How the growing gap in life expectancy may affect retirement benefits and reforms. Mannheim, Germany: National Bureau of Economic Research.

Becker GS (1985). Human capital, effort, and the sexual division of labor. Journal of Labor Economics, 3, S33–S58.

Börsch-Supan A, & Jürges H (2005). The survey of health, aging, and retirement in Europe—Methodology. Mannheim, Germany: Mannheim Research Institute for the Economics of Aging.

Bound J, Schoenbaum M, & Waidmann T (1995). Race and education differences in disability status and labor force attachment in the health and retirement survey. Journal of Human Resources, 30, S227–S267.

Brooks C, & Bolzendahl C (2004). The transformation of U.S. gender role attitudes: Cohort replacement, social-structural change, and ideological learning. Social Science Research, 33, 106–133.

Calasanti TM, & Bailey CA (1991). Gender inequality and the division of household labor in the United States and Sweden: A socialist-feminist approach. Social Problems, 38, 34–53.

Carlson DL, & Lynch JL (2017). Purchases, penalties, and power: The relationship between earnings and housework. Journal of Marriage and Family, 79, 199–224.

Chen Y, & Sloan FA (2015). Explaining disability trends in the U.S. elderly and near-elderly population. Health Services Research, 50, 1528–1549. [PubMed: 25655273]

Crimmins EM, Hayward MD, Hagedorn A, Saito Y, & Brouard N (2009). Change in disability-free life expectancy for Americans 70 years old and older. Demography, 46, 627–646. [PubMed: 19771948]

Crimmins EM, Kim JK, & Solé-Auró A (2011). Gender differences in health: Results from SHARE, ELSA and HRS. The European Journal of Public Health, 21, 81–91. [PubMed: 20237171]

Crimmins EM, & Saito Y (2000). Change in the prevalence of diseases among older Americans: 1984– 1994. Demographic Research, 3.

Davis SN, & Greenstein TN (2004). Cross-national variations in the division of household labor. Journal of Marriage and Family, 66, 1260–1271.

Domingue BW, Belsky DW, Harrati A, Conley D, Weir DR, & Boardman JD (2017). Mortality selection in a genetic sample and implications for association studies. International Journal of Epidemiology, 46, 1285–1294. [PubMed: 28402496]

Edwards JD, Lunsman M, Perkins M, Rebok GW, & Roth DL (2009). Driving cessation and health trajectories in older adults. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 64, 1290–1295.

Fleishman JA, Spector WD, & Altman BM (2002). Impact of differential item functioning on age and gender differences in functional disability. The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences, 57, S275–S284.

Freedman VA, Aykan H, Wolf DA, & Marcotte JE (2004). Disability and home care dynamics among older unmarried Americans. The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences, 59(1), S25–S33.

Freedman VA, Martin LG, & Schoeni RF (2002). Recent trends in disability and functioning among older adults in the United States: A systematic review. The Journal of the American Medical Association, 288, 3137–3146. [PubMed: 12495394]

Freedman VA, Spillman BC, Andreski PM, Cornman JC, Crimmins EM, Kramarow E, … Waidmann TA (2013). Trends in late-life activity limitations in the United States: An update from five national surveys. Demography, 50, 661–671. [PubMed: 23104207]

Fuwa M (2004). Macro-level gender inequality and the division of household labor in 22 countries. American Sociological Review, 69, 751–767.

Geist C, & Tabler J (2018). Somebody has to DUST! Gender, health, and housework in older couples. Journal of Women & Aging, 30, 38–48. [PubMed: 28166470]

Hayutin AM (2007). Graying of the global population. Public Policy and Aging Report, 17(4), 12–17.

Hayward MD, & Sheehan CM (2016). Does the body forget? Adult health, life course dynamics, and social change In Handbook of the life course (pp. 355–368). London, England: Springer.

Henning-Smith C, Shippee T, & Capistrant B (2017). Later-life disability in environmental context: Why living arrangements matter. The Gerontologist. Advance online publication. doi:10.1093/ geront/gnx019

Hybels CF, Pieper CF, Blazer DG, Fillenbaum GG, & Steffens DC (2010). Trajectories of mobility and IADL function in older patients diagnosed with major depression. International Journal of Geriatric Psychiatry, 25, 74–81. [PubMed: 19548209]

Masters RK, Hummer RA, & Powers DA (2012). Educational differences in U.S. adult mortality a cohort perspective. American Sociological Review, 77, 548–572. [PubMed: 25346542]

McRobbie A (2009). The aftermath of feminism: Gender, culture and social change. London, England: SAGE.

Melvin J, Hummer R, Elo I, & Mehta N (2014). Age patterns of racial/ethnic/nativity differences in disability and physical functioning in the United States. Demographic Research, 31, 497–510. [PubMed: 26893587]

Millán-Calenti JC, Tubío J, Pita-Fernández S, González-Abraldes I, Lorenzo T, Fernández-Arruty T, & Maseda A (2010). Prevalence of functional disability in activities of daily living (ADL),

instrumental activities of daily living (IADL) and associated factors, as predictors of morbidity and mortality. Archives of Gerontology and Geriatrics, 50, 306–310. [PubMed: 19520442]

Murtagh KN, & Hubert HB (2004). Gender differences in physical disability among an elderly cohort. American Journal of Public Health, 94, 1406–1411. [PubMed: 15284051]

Nguyen TT, Rist PM, & Glymour MM (2016). Are self-reported neighbourhood characteristics associated with onset of functional limitations in older adults with or without memory impairment? Journal of Epidemiology and Community Health, 70, 1017–1023. [PubMed: 27154180]

Nitsche N, & Grunow D (2016). Housework over the course of relationships: Gender ideology, resources, and the division of housework from a growth curve perspective. Advances in Life Course Research, 29, 80–94.

Palmer RF, Espino DV, Dergance JM, Becho J, & Markides K (2012). The role of physical activity and diabetes status as a moderator: Functional disability among older Mexican Americans. Age and Ageing, 41, 752–758. [PubMed: 23052844]

Percheski C (2008). Opting out? Cohort differences in professional women's employment rates from 1960 to 2005. American Sociological Review, 73, 497–517.

Ryder NB (1965). The cohort as a concept in the study of social change. American Sociological Review, 30, 843–861. [PubMed: 5846306]

Schoeni RF, Martin LG, Andreski PM, & Freedman VA (2005). Persistent and growing socioeconomic disparities in disability among the elderly: 1982–2002. American Journal of Public Health, 95, 2065–2070. [PubMed: 16254235]

Shaw M, Haas S, Krupp L, Clayton A, Langdon D, & Charvet L (2017). Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS) predicts performance on instrumental activities of daily living (P4. 332). Neurology, 88(16 Suppl.), P4. 332.

Sheehan CM, & Tucker-Drob EM (2017). Gendered expectations distort male–female differences in instrumental activities of daily living in later adulthood. The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences. Advance online publication. doi:10.1093/geronb/gbw209

Spector WD, Katz S, Murphy JB, & Fulton JP (1987). The hierarchical relationship between activities of daily living and instrumental activities of daily living. Journal of Chronic Diseases, 40, 481–489. [PubMed: 3597653]

Spillman BC (2004). Changes in elderly disability rates and the implications for health care utilization and cost. Milbank Quarterly, 82, 157–194. [PubMed: 15016247]

Steinfeld E, & Danford GS (1999). Enabling environments: Measuring the impact of environment on disability and rehabilitation. New York, NY: Springer Science & Business Media.

Taylor MG, & Lynch SM (2011). Cohort differences and chronic disease profiles of differential disability trajectories. The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences, 66, 729–738.

Teerawichitchainan B, Knodel J, Loi VM, & Huy VT (2010). The gender division of household labor in Vietnam: Cohort trends and regional variations. Journal of Comparative Family Studies, 41, 57–85.

Thomeer MB, Mudrazija S, & Angel JL (2016). Relationship status and long-term care facility use in later life. The Journals of Gerontology, Series B: Psychological Sciences & Social Sciences, 71, 711–723.

Verbrugge LM, Brown DC, & Zajacova A (2017). Disability rises gradually for a cohort of older Americans. The Journals of Gerontology, Series B, 72, 151–161.

Verbrugge LM, & Jette AM (1994). The disablement process. Social Science & Medicine, 38, 1–14. [PubMed: 8146699]

Vincent GK, & Velkoff VA (2010). The next four decades: The older population in the United States: 2010 to 2050. Washington, DC: U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

White-Means SI, & Rubin RM (2008). Parent caregiving choices of middle-generation blacks and whites in the United States. Journal of Aging and Health, 20, 560–582. [PubMed: 18448685]

Wong R, Michaels-Obregon A, & Palloni A (2017). Cohort profile: The Mexican health and aging study (MHAS). International Journal of Epidemiology, 46, e2. [PubMed: 25626437]

Yong V, Saito Y, & Chan A (2010). Changes in the prevalence of mobility limitations and mobile life expectancy of older adults in Singapore, 1995–2005. Journal of Aging and Health, 22, 120–140. [PubMed: 19934444]

Zajacova A, & Burgard SA (2013). Healthier, wealthier, and wiser: A demonstration of compositional changes in aging cohorts due to selective mortality. Population Research and Policy Review, 32, 311–324. [PubMed: 25075152]
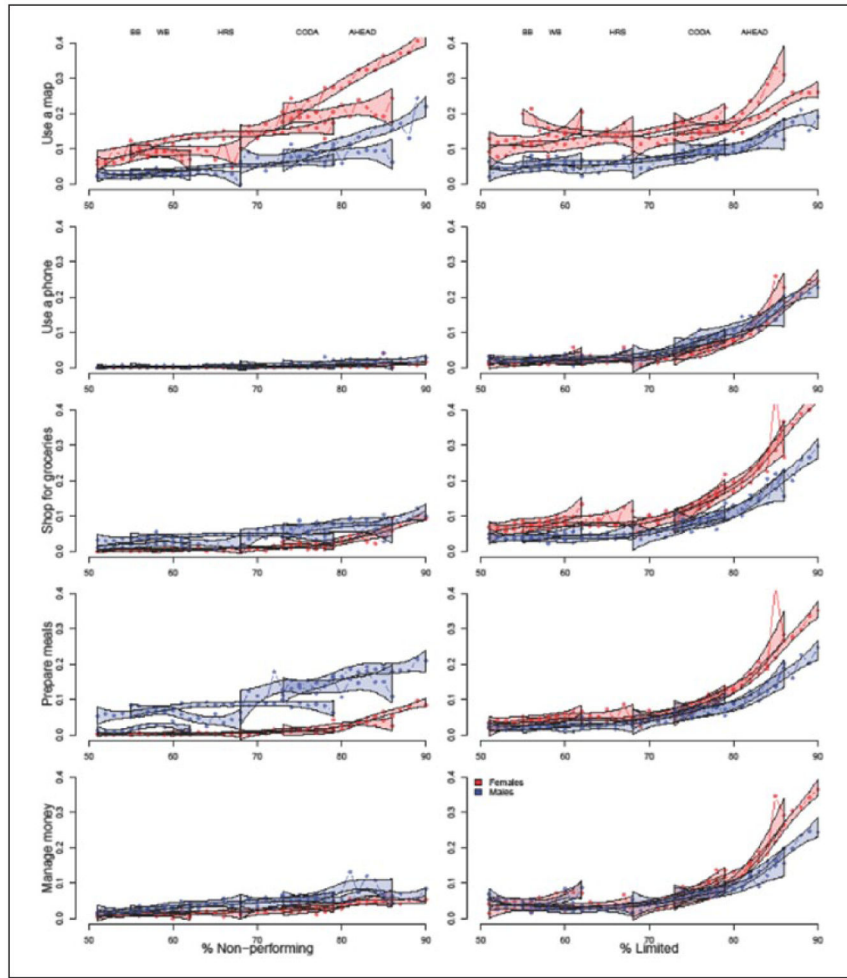
Figure 1.

Age trajectories of IADL nonperformance and health limitations by cohort and gender.

Source. Health and Retirement Study 1995–2010.

Note. The points on the figure represent cohort-specific means by gender for those at a given age. Based on data from each birth cohort, LOESS curves with 95% confidence intervals for reports of nonperformance or limitation as a function of age. AHEAD born before 1924, CODA born 1924–1990, HRS born 1931–1941, WB born 1942–1947, EBB born 1948–1953, MBB 1954–1959. IADL = Instrumental Activities of Daily Living; AHEAD = Asset and Health Dynamics Among the Oldest Old; CODA = Children of the Depression Age; HRS = Health and Retirement Study; WB = War Babies; EBB = Early Baby Boomers; MBB = Mid-Baby Boomers.
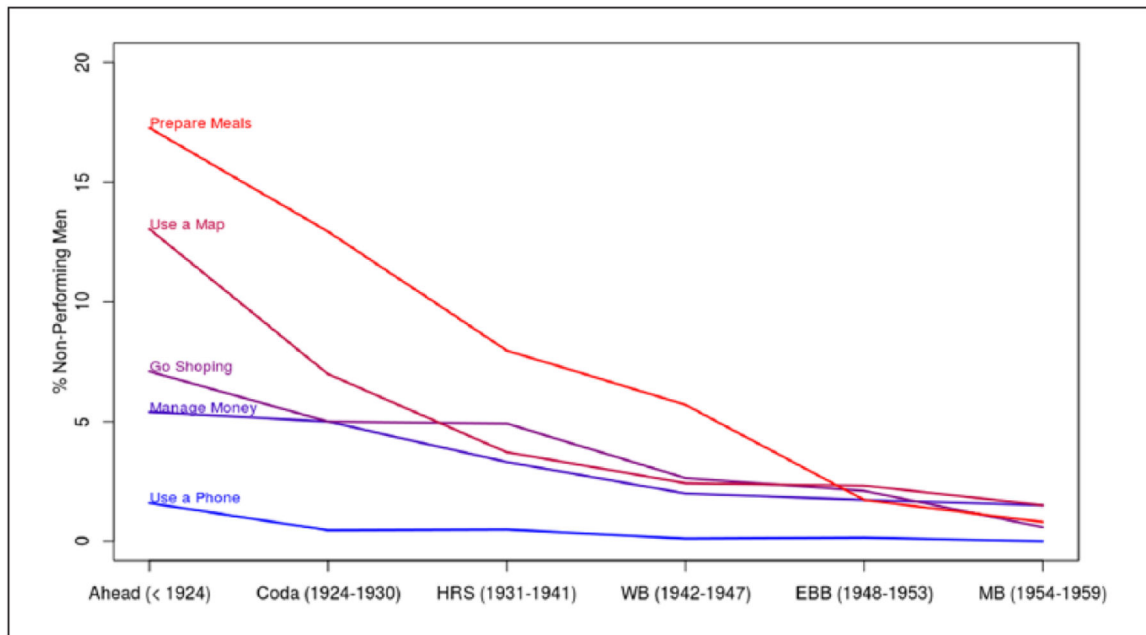
Figure 2.
Percent of men who report nonperformance in each queried IADL by cohort, Health and Retirement Study.

Source. Health and Retirement Study 1995–2010.

Note. Proportion of Men who indicate they "do not do" or are not performing the queried activity by cohort. IADL = Instrumental Activities of Daily Living.

Table 1.

Implication of Reports of Disengagement on IADL Limitations Estimates, by Period and Gender, Health and Retirement Study, 2000–2010.

	Women			Men		
	2000	**2010**	**Change**	**2000**	**2010**	**Change**
Composite measures						
Prevalence						
Any IADL limitation percentage (assume disengaged reports missing)	23.3	27.3	3.9	12.8	17.4	4.7
Any IADL limitation percentage (assume disengaged reports not limited)	28.7	29.9	1.2	22.5	20.4	−2.2
Any IADL limitation percentage (assume disengaged reports limited)	42.2	37.5	−4.7	36.9	28.0	−8.9
Severity (sum of number of limitations)						
Total IADL limitations (assume disengaged reports missing)	0.5	0.5	0.1	0.3	0.4	0.1
Total IADL limitations (assume disengaged reports not limited)	0.5	0.6	0.1	0.4	0.4	0.0
Total IADL limitations (assume disengaged reports limited)	0.8	0.8	0.0	0.7	0.5	−0.1

Source. Health and Retirement Study, 2000 and 2010 Waves. Adults 51+.

Note. More detailed information regarding this table is presented in Supplemental Table 1.

IADL = Instrumental Activities of Daily Living.

Table 2.

Descriptive Statistics of IADL, by Cohort and Gender, Health and Retirement Study, 1995–2010.

	Women						Men					
	AHEAD	CODA	HRS	WB	EBB	MBB	AHEAD	CODA	HRS	WB	EBB	MBB
Year/wave entrance	1995/3	1998/4	1996/3	1998/4	2004/7	2010/10	1995/3	1998/4	1996/3	1998/4	2004/7	2010/10
Average age	79.6	70.3	59.4	52.7	52.8	53.2	78.7	71.1	59.4	53.0	53.1	53.4
Self-reported difficulty using a map												
No	48.7	70.6	67.2	81.5	81.8	80.3	73.9	86.5	89.0	93.9	92.9	91.6
Yes/can't do	20.7	12.4	20.3	9.7	11.4	13.8	13.1	6.5	7.3	3.7	4.8	6.9
Nonperforming	30.7	17.0	12.6	8.8	6.8	5.9	13.0	7.0	3.7	2.4	2.3	1.5
Self-reported difficulty using a telephone												
No	89.7	97.6	98.1	98.4	98.8	97.7	87.4	95.1	96.8	97.5	98.2	97.3
Yes/can't do	9.6	2.4	1.8	1.3	1.1	2.2	11.0	4.5	2.7	2.4	1.7	2.6
Nonperforming	0.8	0.0	0.1	0.4	0.1	0.1	1.6	0.5	0.5	0.1	0.2	0.1
Self-reported difficulty managing money												
No	81.1	95.0	95.4	96.5	94.5	92.6	83.1	90.2	93.8	94.1	94.2	91.9
Yes/can't do	15.6	3.2	3.2	2.6	4.1	6.6	11.5	4.7	2.9	3.7	4.1	6.6
Nonperforming	3.3	1.9	1.5	0.9	1.4	0.8	5.4	5.2	3.3	2.2	1.7	1.5
Self-reported difficulty taking medicine												
No	91.7	97.4	97.9	97.4	97.5	96.9	94.0	97.6	98.4	98.1	97.2	97.2
Yes/can't do	8.3	2.6	2.1	2.6	2.5	3.1	6.0	2.4	1.6	1.9	2.8	2.8
Self-reported difficulty shopping												
No	72.5	90.3	91.8	92.9	93.2	91.5	79.4	89.9	91.3	93.5	95.0	94.4
Yes/can't do	23.2	8.8	7.7	6.8	6.5	8.3	13.5	5.4	3.8	3.9	2.9	5.0
Nonperforming	4.3	0.9	0.5	0.4	0.3	0.2	7.1	4.7	4.9	2.6	2.1	0.6
Self-reported difficulty preparing meals												
No	81.4	93.3	96.0	97.0	96.1	95.4	72.7	83.0	89.5	92.5	96.6	96.4
Yes/can't Do	15.9	5.9	3.8	2.7	3.7	4.5	10.1	4.1	2.5	1.8	1.7	2.8
Nonperforming	2.7	0.8	0.3	0.4	0.2	0.2	17.3	12.9	8.0	5.7	1.7	0.8
N	3,863	1,290	4,543	853	1,275	1,950	2,369	874	3,902	948	1,329	1,851

Source. Health and Retirement Study.

Note. Values in bold are Significantly ($p < .01$) different report of nonperformance than previous cohort as measured by a proportion t test.

AHEAD = Asset and Health Dynamics Among the Oldest Old; COD = Children of the Depression Age; HRS = Health and Retirement Study; WB = War Babies; EBB = Early Baby Boomers; MBB = Mid-Baby Boomers; IADL = Instrumental Activities of Daily Living.

Table 3.

Relative Risk Ratios From Multinomial Regression Models Predicting Responses from IADL Battery Health and Retirement Study, 1995–2010.

	Map		Phone		Money	
	No vs. limited	No vs. nonperforming	No vs. limited	No vs. nonperforming	No vs. limited	No vs. nonperforming
Gender						
Female (Ref.)						
Male	0.35 ***	0.29 ***	1.52 ***	3.27 ***	0.94	1.26
Cohort	0.98	0.85 ***	1.31 ***	0.98	1.56 ***	0.96
Cohort × male	0.99	0.96	1.00	1.00	1.05	1.07
Constant	0.33 ***	0.22 ***	0.00 ***	0.00 ***	0.00 ***	0.03 ***
	Medicine		**Shopping**		**Meals**	
	No vs. limited	No vs. nonperforming	No vs. limited	No vs. nonperforming	No vs. limited	No vs. nonperforming
Gender						
Male	0.79+		0.67 ***	1.10	0.95	5.48 ***
Cohort	1.39 ***		1.29 ***	0.62 ***	1.32 ***	0.58 ***
Cohort × male	1.05		0.98	1.61 ***	0.93 *	1.34
Constant	0.00 ***		0.02 ***	0.03 ***	0.01 ***	0.02 ***
N = 25,047						

Source. Health and Retirement Study. Adults aged 51+.

Note. Models additionally control for age, race/ethnicity, educational attainment, and marital status. Full results presented in Supplemental Table 3. IADL = Instrumental Activities of Daily Living.

[†]
p < .1.

[*]
p < .05.

[**]
p < .01.

[***]
p < .001.