

How Perceptions of Time-Use Deviate from Reality

By Jacob Felson

Americans perceive time as scarcer than in the past, despite greater amounts of free time (Robinson and Godbey, 1997). According to recent research, Americans spend less time at work and more time at leisure than thirty years ago, but believe the opposite. One way of measuring the gap between attitudes about time and behavior is by comparing “stylized” estimates of time use with estimates based on the time-diary approach. Stylized estimates are based on direct time-use questions, such as: “How much time did you spend yesterday cleaning house?” Time diaries involve respondents giving hour-by-hour reports of their activities—typically for a single day--which are then coded into general categories.

Many scholars have concluded that the time diary approach yields more accurate estimates of time spent on most activities than stylized estimates (e.g. Robinson, 1985; Juster and Stafford, 1991; Marini and Shelton, 1993).¹ With few exceptions, studies have concluded that time diaries come closer to describing the way people actually spend their time. This makes sense for a number of reasons. By retracing their steps through the day, time diaries jog the memory of respondents to better remember each episode of activity. Also, respondents filling in time diaries are probably more likely to focus on the requested day, rather than revert to their “typical” time use patterns, or norms or stereotypes about time use (Robinson and Godbey, 1997: p.59).

Researchers have found that gaps for time spent in paid work have widened each decade from 1965 to 1985 (Juster and Stafford, 1991; Robinson and Bostrum, 1994). Respondents exaggerate paid work and under-report total free time (Robinson and Godbey, 1997). Robinson and Bostrum attributed the rising gap in time spent in paid work to a rise in service jobs without strict hourly schedules. In contrast, Robinson and Godbey (1997) cite qualitative evidence suggesting increased blending of work and leisure time. People feel more rushed nowadays because they spend their free time in shorter, less leisurely intervals. The authors did not test this idea statistically.

As the estimates between stylized and time diary estimates have diverged, the gap has gained import, both methodologically and substantively. Methodologically, the wide discrepancy reveals severe problems with widely used stylized estimates, particularly for labor market researchers (Juster and Stafford, 1991). Substantively, the discrepancy between time diary and stylized questions is evidence of an increasing gap between perceptions and behavior, which may have social consequences. The fact that the gap has *increased* over the last thirty years strongly suggests that perceptions of time have changed. Increased gaps reflect the current cultural notion that time is scarcer than ever (see Robinson and Godbey, 1997).

This study focuses on how gaps between different measures of time-use vary according to respondent characteristics, such as patterns of time use. By using a dataset that combines stylized and time diary estimates for *the same day* among the *same respondents*, we are able to take a more detailed look at how misestimates of time-use

¹ Exceptions include criminal, sexual and bathroom activities, and to a lesser extent, telephone use (Juster and Stafford, 1991). Researchers obviously need to ask specifically about highly sensitive, private behaviors like sex to get non-zero estimates.

occur than previous studies. We look at how respondents' patterns of time use affect the extent to which perceptions reflect behavior. As such, we take stylized estimates seriously as rough indicators of social perceptions—not merely as flawed estimates of actual time use gathered at lower expense.

We focus on the gaps between stylized and time diary estimates for time spent in three activities: paid work, housework, and television. These activities represent three of the four categories in Robinson and Godbey's (1997: p. 12) taxonomy of time-use in America: "contracted time" (work), "committed time" (housework) and "free time" (television). As such, the activities examined in this study constitute the majority of the waking hours of many Americans.

Literature Review

Gaps typically reflect exaggeration of time spent in various activities, with the exception of free time, which is typically underreported. Some early evidence of exaggeration comes from Robinson (1985), who compared the means of time spent based on weekly estimates with one-day time diaries. By weighting each diary day equally and aggregating up to a seven-day week, Robinson was able to compare diary and stylized estimates for a "typical" week given by the same respondents.² Findings indicate that respondents generally exaggerated the amount of time spent on activities in their stylized weekly estimates, compared to their time diary reports. We know that general weekly estimates are exaggerated because the means total up to more than the total 168 hours in a week, while the design of the time diary forces activities to add up appropriately.

Few studies have examined how and why people misestimate time spent on various activities differently. Theories have been suggested, but not tested. Studies that include comparisons of diary and stylized estimates are primarily focused on issues unrelated to actually *explaining* the gaps found between modes. Some studies aim to demonstrate the superior validity of the time-use diary method, compared to stylized estimates (e.g. Robinson, 1985; Juster and Stafford, 1991; Marini and Shelton, 1993). These methodologically oriented studies have established the substantial boost in validity gained from employing diary, as opposed to direct ("stylized") question modes.

For example, based on the size of the gaps between these two modes, Robinson (1985) argues that general estimates are simply unacceptable sources of data for estimating actual time-use. Furthermore, Juster and Stafford (1991) assert that the use of stylized estimates poses major problems in the field of labor economics, potentially biasing effect sizes, such as the relationship between having young children at home on hours worked. This is true only to the extent that gaps vary *systematically* by relevant respondent characteristics. Few studies have demonstrated the extent to which gaps vary according to any substantial number of respondent characteristics.

² Extrapolating from a diary day to a week for comparison with stylized estimates poses a problem if we want to use the gap as a dependent variable in a regression equation. This is because gaps between aggregated diary and stylized estimates will arise both from respondent perceptual errors and *actual* behavioral differences. In other words, in a model predicting the size of the gap, perceptual error is confounded with the relative "typicality" of the diary day, for which we could only partly control. In addition, "typical" days are hard to define in a society in which 40% of respondents report their randomly chosen diary day as "atypical" (Robinson and Godbey, 1997).

Some studies have demonstrated how gaps differ according to basic demographic characteristics. Typically, researchers compare estimates culled from time diary data with stylized estimates from other datasets.³ For example, Robinson and Bostrum (1994) looked at how workweek estimates from the 1985 Current Population Survey (CPS) differed from estimates based on a time-diary study in the same year. They found that the gap between estimates was linearly related to number of hours paid work in the diary. People working longer hours were more likely to exaggerate their work hours, while people working few hours were more likely to underestimate their hours. The authors also found an interaction between gender and numbers of hours worked.⁴ While women who worked few hours tended to give relatively accurate stylized estimates, women working longer hours were more likely to exaggerate their hours than men.

Robinson and Godbey (1997: p. 91) suggest this interaction results from traditional gender roles. Since women tend to have more experience in part-time jobs, their time use in these situations is more familiar and thus more easily estimated. By contrast, women working many hours may exaggerate more because their reference point is “no hours” under a traditional gender roles regime. Since men have more experience with longer hours, they have less trouble estimating their work hours. This explanation remains an untested supposition.

Previous studies suggest that misestimation of time spent in other activities also varies by gender. Robinson (1985) presented tables revealing that women appear to exaggerate television viewing and “house chores,” as well as paid work, more than men (1985). If gender roles played a large role, we might expect women to provide *better* estimates of housework, since they are more familiar with these tasks relative to men.

Juster and Stafford (1985) find a more extreme correlation between time use gaps and gender in estimating time spent on housework. Comparing stylized estimates from a study by Morgan et al (1996) with diaries from a 1965-66 survey, they found that housework hours are *overreported* by women, but *underreported* by men. The authors found evidence that the substantial gap results from men’s definition of housework, which includes fewer activities than women’s.

In fact, the gender gap in misreporting of housework disappeared in recent years with more refined questions. Bianchi et al (2000) presents comparisons of figures from the 1992-94 National Survey of Families and Households (NSFH2) with a 1995 time diary study. The NSFH2 broke down housework into nine relatively detailed questions. As a result, the gender gap was zero; both men and women exaggerated their housework by about fifty percent, based on a comparison between means from the time diary survey and NSFH2.

Furthermore, Robinson (1985) showed that time use gaps can be partly explained by considering background or “secondary activities” recorded in the time diary. In most time diaries, after respondents report primary activities, they are probed for further

³ Comparing estimates from different datasets is adequate for looking at basic demographic differences-- barring substantial differences between samples in coverage and non-response. But looking at more specific respondent characteristics with respect to patterns of time usage requires comparison of questions on the *same time diary survey*.

⁴ Robinson and Bostrum did not actually run regressions demonstrating this interaction, since they were dealing with two different samples. Thus, the term interaction here is used loosely.

activities they may have been doing simultaneously. These activities are coded as secondary and tertiary activities.⁵

When asked to give general estimates, Robinson found that respondents tend to count some of these secondary activities, leading to “double counting.” For example, if a respondent did housework while watching television for two hours, they might include those two hours in their answers to questions about housework *and* television. But Robinson concluded that only part of the gaps could be explained by so-called “double counting” time spent in simultaneous activities.

Other studies have compared stylized with time diary estimates, but primarily in the context of unrelated substantive issues, such as the gender division of household labor (e.g. Bianchi, et al, 2000) and church attendance frequency (Presser and Stinson, 1998). Presser and Stinson (1998) introduced time-diary estimates of church attendance as a validity check on the accuracy of stylized estimates and count data.

The recent debate about measuring church attendance sheds light on explaining gaps. Standard questions asking respondents whether they attended church last week consistently yield estimates of about 38% to 40% attending church on a given Sunday (Smith, 1998). Estimates based on diaries for Sunday yield declining estimates that now stand at about 25%. Researchers conclude that an increasing percentage of people are saying they went to church on Sunday when they did not. Marler and Hadaway (1999) find evidence suggesting that many of the people who did not go to church but say they do are in fact more frequent attendees. Many people interpret the standard question “Did you have a chance to go to church last Sunday?” as a question about the “typical” Sunday.

This explanation for the church attendance gaps is similar to explanations offered for more frequent activities. When respondents are asked about a typical amount of paid work, they tend to think of the *typical day in which a non-zero amount of paid work was performed*. In other words, respondents tend to ignore days when the activity simply was not performed, thereby inflating estimates. This problem is minimized when interviewers ask about “yesterday,” rather than “a typical day.”

Time-Deepening and Increased Time-use Mode Gaps

Robinson and Godbey (1997) suggest that Americans’ greater misperception of time is related to increased blending of activities. They argue that the reason we feel so harried—despite our increased leisure time—is our propensity to do more things simultaneously, and to blend multiple activities at a frenetic pace. For example, when we go back and forth between personal email and work, we may *feel* nearly as tired at the end of the day, as if we had worked without ceasing. Multi-tasking may lead us to exaggerate our work hours and underreport our leisure time. We do not feel the full effect of free time unless it is spent in large quantities, according to the theory.

⁵ It might be more accurate to say that a person alternates back and forth very quickly between several activities. Few people can really do two activities at *exactly* the same time.

The Present Study

I base my hypotheses for this study on Robinson and Godbey's (1997) suppositions and on evidence from previous research. Unlike previous research, this study looks at misestimation more generally, rather than over- or under-reporting. Based on prior research, I predict that people who spend more time on an activity will generally have a harder time estimating it. Longer activities tend to come with more breaks. Activities we do habitually but for shorter periods of time are easier to estimate, since they are more likely to be uninterrupted. Thus, I predict larger time use gaps for respondents whose activities are broken down into more discrete intervals of time. More intervals of time imply more breaks in between tasks that can be inadvertently "counted" as work.

I also expect people who do more things at once will have higher gaps. Simultaneous activities can be conceptualized more accurately as activities which respondents alternate between very rapidly. Thus, we can see secondary activities in a time diary as quantitatively, rather than qualitatively, different from activities listed in different time slots. At some unknown level of quickness, alternating activities become perceived as "simultaneous" as opposed to merely rapidly alternating.

Other kinds of time patterns should help explain gaps. People who allot similar periods of time to each activity will give more accurate stylized estimates and lower gaps. Conversely, people who spend vastly different amounts of time on different activities will have higher gaps. Five-minute intervals of one activity may get lost amidst longer activities spent doing other activities. This idea follows from suggestions by Robinson and Godbey (1997). For instance, short intervals of free time (say on the internet) may get lost within a long day at work, leaving people working less but feeling they've worked more.

If Robinson and Godbey (1997) are correct about the way conceptions of time have changed, older cohorts may have fewer gaps for reasons stated above. Age will be related to gaps only if people's patterns and perceptions of time are actually determined during an impressionable period in their youth, rather than change over time as society changes. Age would negatively correlate with gaps if Robinson and Godbey are correct, *and* cohort effects exist. Otherwise, age could be positively correlated with gaps, since people's memories deteriorate with age.

Moreover, Robinson and Godbey's work implies that people who subjectively report feeling more "rushed" might have bigger gaps, especially for work activities. People who feel rushed may be more likely to feel they're doing a lot of work (paid or in the house), when they are doing a moderate amount, but a harried pace of alternating activities leaves them drained.

Gaps for the three activities examined here should correlate with each other. Gaps for activities that are uncorrelated with other gaps suggest explanations peculiar to the activity, not to perceptions of time generally. Gaps of different activities that inter-correlate suggest that perceptions of time, influenced by patterns of time generally, generate most misestimation.

A variety of other patterns of time use reporting should be associated with estimation gaps. People who are more precise in their time diaries should have misreport less in stylized estimates. Thus, respondents who give more beginning and ending times

rather than simple numbers of hours should give fewer gaps. Respondents who give more precise beginning and ending times for activities should have lower gaps. Informal activities generally do not begin and end at the hour and half-hour, even though we have a tendency to report these times. Thus, respondents who report beginning and ending times at *other* times (e.g. 12:05 rather than 12:00 or 12:30) probably have fewer gaps.

A variety of demographic characteristics should affect how fragmented and frenetic time use is, as indicated by the measures described above. Parents with young children at home—especially single parents—should lead more frenetic lives, and misestimate their time-use more. Higher SES persons might do more multi-tasking, and thus would give worse estimates. People living in large cities might give worse estimates for similar reasons. Conversely, residents of the ‘Old South’ might have smaller estimates if they live life with a less frenetic pace. On the other hand, a less precise sense of time might lead Southerners to higher gaps.

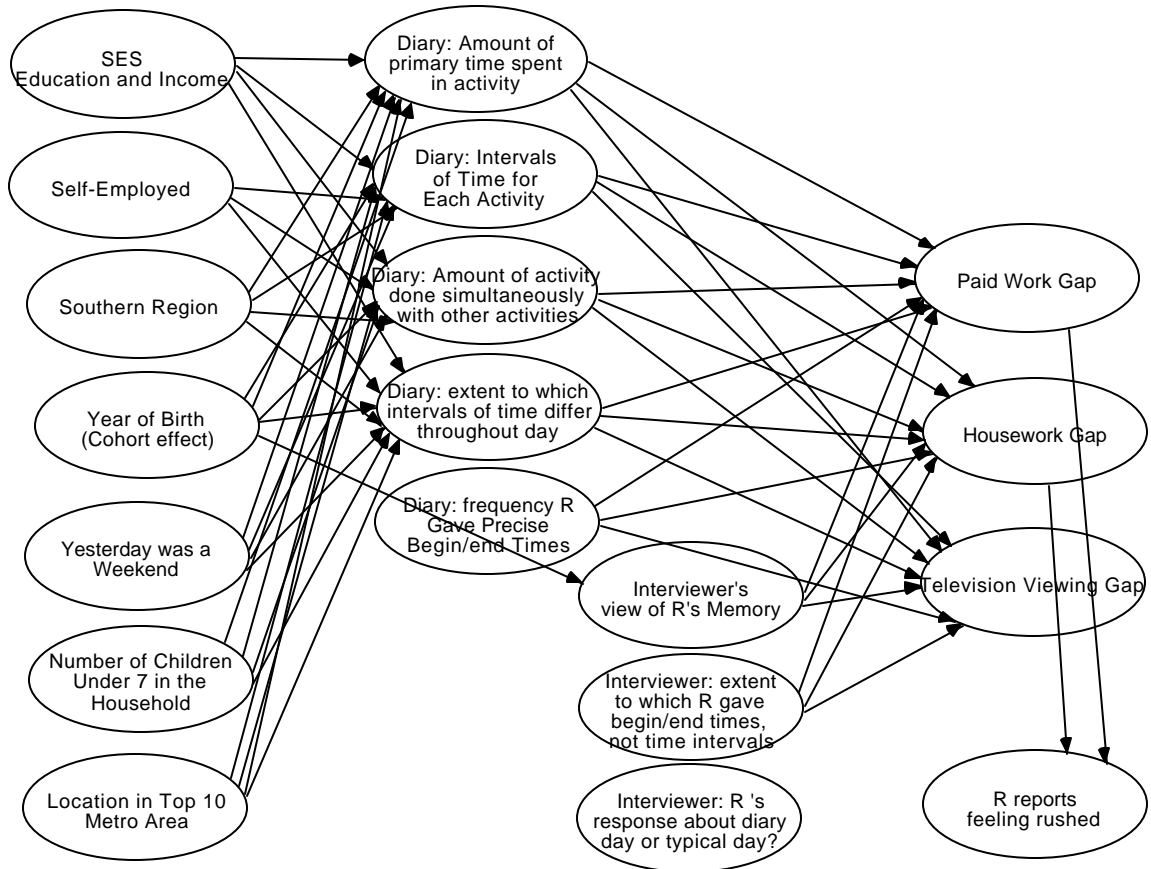
The diagram presented below provides a more formal conceptualization of the hypotheses presented above. Patterns of time mediate the relationship between characteristics of the household (and the particular day) and gaps. Patterns of time may also act independently as exogenous variables.

The bottom two variables presented in the middle part of the diagram should also influence time-use estimates independently. Finally, the interviewer’s perception of the quality of recall is also important, in part because it may suppress the effects of higher SES and age on gaps. More highly educated persons may have better memories and give better gaps for this reason. But more educated persons may also have more “fragmented” lives with more simultaneous activities, leading them to have larger gaps. Older people might hew to earlier time use patterns, leading them to less frenetic lifestyle, and lower gaps.⁶ But they may also have worse memories, leading them to give worse estimates and have higher gaps.

The final indicator presented in the conceptual diagram (“R reports feeling rushed”) is suggested by Robinson and Godbey’s (1997) qualitative discussions. Respondents feel more rushed because their perceptions of paid and house work are skewed, since they are fragmented and blended. This indicator is a part of the model that will remain untested in the subsequent analysis.

⁶ The effect of age is complicated. People in mid-life tend to be busiest, so the relationship between age and amount of leisure time, for example, is curvilinear. The small dataset used in the present analysis does not have sufficient power to examine highly collinear quadratic effects.

Figure 1: Conceptual Diagram



Data

The dataset is from the Family Interaction, Social Capital, and Trends in Time Use (FISCT) survey (Robinson, Bianchi and Presser, 2001). The University of Maryland Survey Research Center conducted FISCT as a random-digit dial telephone survey of 1,151 respondents between March 1998 and December 1999. The survey was conducted to get a broad range of time use throughout the year to minimize the impact of seasonal variation on estimates. The response rate was 55.5% (Bianchi, Robinson and Sayer, 2001). Time-diaries for the previous day were collected from all respondents.

This study uses a sub-sample (N=276) of FISCT⁷ which contains standard time use questions for “yesterday,” the day in which respondents are subsequently asked to recall activities in diary form. This is the only nationally representative survey that includes both time diaries and stylized estimate questions about the *same day*. This unique combination allows us to attribute discrepancies to exclusively misestimation of time.

This advantage comes with some disadvantages. First, there may be some consistency effects minimizing gaps that would otherwise be much larger. Respondents may remember their answers to previous questions about how long they worked or watched television, and apply this to time diary responses. Respondents may be steered in the direction of responses they gave to previous stylized questions. To the extent that consistency effects are invariant across groups, this issue should not bias effects. But the relatively small gaps and the small sample size will limit power. Gaps reported in this study are very conservative estimates, the magnitudes of which should not be compared with gaps reported elsewhere between different surveys.

Previous time diary designs have involved prospective “tomorrow” mail-back designs, as opposed to yesterday “telephone designs.” Intuitively, it seems that respondents would have more time to think and give more accurate, detailed lists of activities in mail-back paper-and-pencil diaries than telephone surveys. However, several studies have shown estimates culled from recall diaries and mail-back diaries are comparable (Scheuch, 1972; Robinson, 1977).

In the data used here, interviewers generally reported that respondents had “little difficulty” remembering their “diary day.” Only 5% of respondents in the sub-sample used here were considered by interviewers to have “great difficulty” with recall.

Methods

We employ three sets of OLS regressions for this analysis. All analyses are weighted with a variable provided in the FISCT dataset. This weight adjusts for survey

⁷ FISCT consisted of several ballots randomly-assigned to respondents, including one ballot with questions about activities in a typical week. For some purposes, this ballot gives analysis more power, due to a larger sample size. However, there are problems with comparing yesterday times diaries with typical week estimates. In the present analysis, we need to assume that diary estimates are closer to reality than stylized estimates. We cannot make this assumption if we compare weekly estimates with extrapolated *daily* estimates.

design and post-stratification.⁸ The dependent variables are “gaps” based on the difference between stylized estimates, and time diary estimates of total time spent in each activity (*absolute value* of Stylized Estimate-Time Diary Estimate).⁹ Gaps were trimmed at the 95th percentile for housework and TV viewing, and at the 99th percentile for paid work so that analyses would not be affected by extreme outliers.¹⁰

The estimation gaps used here are analogous to change scores. There is debate in the methodological literature about the virtues of the *change score method* versus the so-called *regressor variable method* (Allison, 1990). In this analysis, we decided to employ *the change score method* by using gaps as the dependent variables, rather than regress stylized estimates on time diary estimates. This is because we cannot be certain that the time diary estimate is measured without error. In the typical *regressor variable method*, measurement error in the time diary estimate can easily cause other independent variables to be misspecified. This sort of misspecification error is not apparent in change score models, which effectively constrain the time diary coefficient to one (Allison, 1990). There is a dependent variable in the analysis for three activities: paid work, housework and television viewing. The stylized estimates for paid work are based on answers to the question, “How much time did you spend yesterday working for pay?”

Stylized estimates of housework are based on tallying up separate stylized respondent reports of time spent yesterday in six activities: preparing meals; washing dishes and cleaning up after meals; cleaning house; outdoor and household maintenance tasks; shopping for groceries and other household goods; washing, ironing and mending; and paying bills and keeping financial records.¹¹ Time diary use estimates are likewise tallied and categorized.

The third dependent variable is an estimation gap for television viewing. The stylized question about television is: “How much time did you spend yesterday watching TV?” The diary estimates are tallied likewise.

Independent Variables

The first set of variables includes respondent characteristics and a dummy variable indicating the day of the week was a weekend. Education was measured as the number of years of schooling, capped at 18 for persons with graduate or professional degrees. There were no missing cases for education or the weekend dummy.

⁸ In RDD samples, phone number—not each household—had an equal probability of selection. The survey included a question about the number of non-business telephone numbers for the household. The weight, in part, adjusts for houses that had a greater chance of selection. Post-stratification weighting included adjustments to correct for discrepancies between the sample distribution sex, education, race, census region, and 1996 Current Population Survey distributions. To avoid excess variance of the weights, a small number of cases less than two or greater than 5 were “trimmed” (Bianchi, Robinson and Sayer, 2001).

⁹ In analyses not presented, we experimented with multinomial logistic models with three categories: overestimation, accurate (within 30 minutes) and underestimation. These analyses proved difficult because of the low number of degrees freedom with sample size of 276.

¹⁰ See Bianchi et al, 2000 for rationale with time use estimates.

¹¹ Other things equal, housework gaps should be lower because respondents are asked about *specific* activities generally. On the other hand, asking about specific paid work activities is more problematic and less necessary. The definitions of paid work and television viewing are less ambiguous than the definition of housework. So one can argue that more categories are required to standardize interpretation.

Household income was recoded by the survey researchers from a series of questions into categories from (1) \$12,000 or less to (7) \$100,000 or more. The nineteen percent missing cases for the subsample used in this analysis were coded at the mean. A dummy was coded one for missing on income. This variable was not a significant predictor, and was taken out of the analysis.

Dummy variables were coded one for answers to questions about: self-employment, and Southern region. Respondents were asked their age indirectly by year of birth. Number of children under 7 was coded as an interval variable to represent answers to, “How many [children] are 6 or younger?”

Respondents were coded one if they lived in a top 10 metropolitan area, as defined by whether the geographic referent of their area code was included among the top 10 metro areas enumerated by the 2000 Census.

A number of mediating variables were tallied from raw diary data. Refer to the conceptual diagram presented above for the names of these variables. All variables presaged with “Diary” are based on calculations based on raw diary data.

Read from top to bottom in the middle section of variables.

- (1) Amount of primary time spent on an activity is the sum of the time reported first to for each activity (e.g. paid work, housework, or TV viewing).
- (2) Intervals of time for each activity is the sum of the total *number* of times the activity is listed, either as a primary, secondary or tertiary activity.¹²
- (3) Amount of activity done simultaneously with other activities is the sum of minutes reported doing the activity, whether primary, secondary or tertiary.
- (4) The extent to which intervals of time (for all activities recorded) differ throughout the day is the *standard deviation* of all the time intervals in the diary for a given respondent. If all activities took 30 minutes, this indicator would equal zero. If some activities took 5 minutes, and others took several hours, this indicator would be high, relatively speaking.
- (5) The frequency respondent gave precise beginning and ending times is a tally of the number of times respondents gave times *other than* hour or half-hour (e.g. 12:15 rather than 12:00 or 12:30).
- (6) Interviewers were asked to record their perceptions about the extent of difficulty in respondents remembering their diary day on a four point scale from ‘great difficulty’ to ‘no difficulty’.
- (7) The extent to which respondent gave beginning and ending times, not time intervals is also based on an interviewer question. This indicator is based on the following question: “What time did you finish___?” is sometimes answered directly with an end time; other times it is answered indirectly with length of time the activity lasted. In this interview, did the respondent: (1) always give the time activities ended; (2) usually give end times, but sometimes activity lengths; (3)

¹² See Diary Questions in Appendix A. Primary activities are activities reported first. Secondary and tertiary activities were culled from responses to the question: At any time while you were [REPEAT ACTIVITY] did you do anything else? Up to two activities were accepted for this question (secondary and tertiary).

- Sometimes give end times, but usually activity lengths; (4) Always give activity length, not end times.
- (8) The interviewer's perception about whether the respondent based their answers on the diary day or on a typical day was based on the following question: "Did the respondent's answers to the diary day make you think he/she was thinking: (1) only about the diary day, (2) mainly about the diary day, (3) equally about the diary day and how he/she usually spends time, (4) mainly about how he/she usually spends time. Responses were dichotomized into a dummy variable coded 1 for interviewer perception only or mainly of "typical day" responses.

Finally, the measure of the extent to which a respondent feels rushed is based on the following question: "Would you say you always feel rushed even to do the things you have to do, only sometimes rushed, or almost never feel rushed?" This variable was dichotomized into a dummy variable coded '1' for "always" rushed, and '0' for other responses.

Results

Past research has shown that some respondents "double count" secondary activities. If this were the case, then including secondary activities in calculating the gaps should result in gaps with lower means. But as the descriptive statistics in Table 1 show, this is not the case; whether or not we include secondary minutes in the difference score makes little difference on the mean gap.¹³ This may be due to the wording of the question, as seen in Appendix A. Respondents are asked whether they did anything else *at any time* during the primary activity. Secondary activities are coded as taking *the same amount of time* as the primary entry, when in fact these activities could be much shorter. Thus, there are forces going in opposite directions, leaving the overall means about the same.

Other descriptive statistics are important for the analysis. We find the rather obvious fact that paid work is more structured than housework or TV viewing. People tend to do paid work exclusively in single intervals of time during the day. By contrast, TV is an activity that blends with other tasks in people's day. On average, about 29% of the 2.8 hours of television watched on the diary day provides a background for another activity. Of the two hours of average housework on the diary day, only 15% of is background for another activity. But housework is done in slightly more frequent intervals than television.

¹³ In regressions not shown, we regressed the gap *excluding* secondary minutes on secondary minutes, and found that secondary minutes was *not* a significant predictor of the gaps. This also provides support for the idea that people may not double count. Keep in mind that in the regressions shown in Tables 2 through 4, the dependent variable is the gap *including* secondary minutes. Thus, regressing secondary minutes on this gap is not an indicator of whether people double count, but whether multi-tasking causes them to misestimate.

{Table 1 here.}

Caution is required in interpreting the estimates for time spent in primary time and secondary time. It is possible that these effects may be statistical artifacts. For example, secondary paid time could be highly correlated with the gap, because few people do secondary paid work. Thus, adding in the estimate for secondary paid work might influence the dependent variable to such an extent that secondary paid work would automatically correlate highly with it.

We tried a different way of testing our main hypothesis in order to try to avoid statistical artifacts. We regressed gaps *excluding* secondary minutes on number of secondary minutes in an activity. We also regressed gaps *excluding* secondary minutes on the percentage of total minutes in an activity that were secondary, rather than primary. The variables of these alternative designs have distinct components, so no statistical artifacts should be present. We found smaller, but significant effect sizes for housework and TV viewing, but no effect for paid work. These results make sense, since paid work is more structured than the other activities. The amount of paid work that is secondary is small and varies less between respondents. In general, these results provide (weaker) support for our hypotheses.

We also considered the possibility that our results were heavily influenced by people who reported doing none of any of the three activities. Zeros predicting zeros in this analysis could cause misspecification. To test for this possibility, we excluded people who gave zero responses to both the diary entry and stylized estimates. Respondents giving zero responses to only one of the estimates were left in, since their time use remains ambiguous. Running the final models in this way, we found very similar effects.

Paid Work

Table 2 shows mixed support for our hypotheses. Higher income respondents show lower, not higher gaps, counter to our hypothesis. Why does income behave counter to predictions? Lower SES respondents may be more prone to socially desirable answers. Lower SES persons may have been more likely to give standard stylized estimates, such as an 8 hour work day, and thus have higher gaps.

Age shows a rather strong negative correlation with gaps. As people age, they may be more likely to settle into job routines they are very familiar with, and find easy to estimate. On the other hand, the effect of age could support our hypothesis, that earlier cohorts are more likely to have a less frenetic lifestyle, and thus less likely to misestimate. However, our measures of “frenetic lifestyle” that survive multivariate analysis do not explain away the effect of age.

Paid intervals do not survive multivariate analysis because it is so highly correlated with the amount of time at work. Consistent with some previous research, we find that people who work longer hours give less accurate work estimates.¹⁴ Secondary paid time works in the same way. One might conclude that a single interpretation is

¹⁴ Previous research showed that people working more hours were more likely to *exaggerate*. Our finding thus remains consistent with, but does not exactly replicate this finding.

sufficient to explain the positive effects of primary and secondary paid time. People who work fewer hours naturally find it easier to estimate them, since remembering a shorter period of time is generally easier.

But the effect of secondary paid time is considerably stronger than the effect of primary paid time. This may support our hypothesis that people who alternate activities such that they report a greater number of simultaneous activities will give worse estimates. If a person watches TV and answers personal phone calls while doing work all day, he may have great difficulty estimating his time.

Self-employed people are significantly more likely to misestimate their time. This conforms to the general idea that less structured time is harder to estimate. People who work for others generally conform to standard rules about leaving a workplace that are ingrained in their mind. Self-employed people set their own hours and can generally come and go as they please. Self-employed people probably give worse estimates of their work hours because they do more non-work activities during regular work hours. We are not able to explain away this effect after controlling for time use patterns, but our measures are rough. The most plausible explanation for the effect of self-employed on the paid work gap remains a pattern of time that is more “blended.”

Housework and TV-viewing

We find support for our hypothesis that more young kids make estimation of activities in the home more difficult. The number of children in the house is strongly related to both housework and TV-viewing gaps. This effect strongly suggests that alternating activities more frequently leads to worse estimates. Parents of young children are probably the most likely adults to be interrupted from their chores around the house, or their favorite television programs.

Education and income are related in the opposite direction predicted. More educated people might have better memories of their stylized estimates. Thus, they might be more susceptible to consistency effects, lowering the gaps. “Good memory” perceptions by the interviewer of their diary day may be only weakly related to this consistency effect.

The standard deviation of time intervals survives multivariate analysis in predicting housework gaps, but not television gaps. Higher standard deviations indicate higher numbers of short activities amidst some much longer activities. Housework is most likely to be one of these short activities. Thus, individuals who report more of these short intervals of housework in the diary, have more trouble adding them up quickly to get an estimate. Here, the number of intervals is not the key. Rather, greater misestimation occurs when smaller intervals of housework are mixed with longer intervals of doing other things. The greater number of small intervals involved, the more difficult it will be for respondents to remember them *and* sum them up to the stylized estimate correctly.

Housework and TV-viewing gaps are correlated at a moderately high level. In analyses not shown, some evidence was found that people who over-estimate housework underestimate TV-viewing. Since paid work tends to take place away from home, time spent on paid work is less likely confused with other activities.

Of course, geography as measured here did not make a difference. Questions about interviewer perceptions were generally weak predictors. Finally, estimation of housework and television viewing was no less accurate on weekends, suggesting that the lack of structure on these days is not an issue here. Predictably, the effect of weekend on paid work gap disappears when we control for the number of hours worked.

{Table 2 here}

Conclusion

Overall, support for the conceptual model is mixed. Certain measures proved somewhat problematic; extremely high correlations may suggest statistical problems. The strongest, less ambiguous support came from ideas about the patterns of households with young children, and self-employed workers. These are two cases where frenetic lifestyles can cause greater misestimation of time. Other effects are open to a variety of interpretations.

This work sets the stage for more rigorous designs examining how time use patterns affect perceptions of time spent. Researchers can take stylized estimates seriously as indicators of perceptions, rather than simply flawed measures to be discarded when possible. Researchers could look for more rigorous measures of time use patterns. For example, the number of small intervals of a particular activity (e.g. under twenty minutes) during the day may be a better measure of how hectic life seems to respondents. Also, researchers could examine how more specific sequences of activities affect perceptions of time spent. Models could include the frequency of change in location of other persons present.

This paper also sets the stage for further analysis of over-time trends. Robinson and Godbey (1997) present quantitative evidence indicating that estimation gaps for work and leisure have increased since 1965. Higher gaps between perception and reality reflect a society with objectively more leisure time, but subjectively less. Increased gaps are presumably related to a societal ethos that emphasizes the scarcity of time.

Robinson and Godbey suggest this ethos stems from increased blending of different activities. When leisure is mixed with work, breaks do not feel as leisurely or leave people as well rested. If a person takes an hour of leisure at the end of the day, rather than six ten-minute breaks, the argument goes, he or she will feel better rested. The whole feels greater than the sum of its parts.

The next step is to quantitatively test this idea. Do people report more frenetic blending of activities in their time diaries over time? Robinson and Godbey (1997) did not specifically address this issue. If this is the case, researchers could perform a decomposition analysis with Firebaugh's (1997) methods. What proportion of the enlarged gaps is due to more blending of activities, variously measured? What proportion is due to less standardization of daily labors, more 'atypical' days? Analysis of change over time would obviously require better measures of activity blending than those used here. Also, decomposition analysis would involve comparing weekly estimates with aggregated diary day estimates. While methodological problems may be significant, explaining the divergence between time perception and reality would be an important contribution.

References

- Allison, Paul. 1990. "Changes Scores as Dependent Variables in Regression Analysis." *Sociological Methodology* 93-114.
- Bianchi, Suzanne, John Robinson and Liana Sayer. 2001. *Family Interaction, Social Capital and Trends in Time Use Project Report*. College Park, MD: University of Maryland Press.
- Bianchi, Suzanne, Melissa Milkie, Liana Sayer, and John Robinson. 2000. "Is Anyone Doing the Housework? Trends in the Gender Division of Household Labor?" *Social Forces*, 79(1): 191-228.
- Firebaugh, Glenn. 1997. *Analyzing Repeated Surveys*. Thousand Oaks, CA: Sage Publications.
- Juster, F. Thomas and Frank Stafford. "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement." *Journal of Economic Literature* 29: (June) 471-522.
- Marini, Margaret Mooney, and Beth Anne Shelton. 1993. "Measuring Household Work: Recent Experience in the United States." *Social Science Research* 22: 361-382.
- Marler, PL and Hadaway CK. 1999. "Testing the Attendance Gap in a Conservative Church." *Sociology of Religion* 60: 2 175-186.
- Morgan, J.N., Sirageldin, I., and Baerwaldt, N. 1966. *Productive Americans*. Ann Arbor, MI: Institute for Social Research, The University of Michigan.
- Robinson, John and Ann Bostrum. 1994. "The Overestimated Workweek? What Time Diary Measures Suggest." *Monthly Labor Review*, August, 11-23.
- Robinson, John and Geoffrey Godbey. 1997. *Time for Life*. University Park, PA: The Pennsylvania State University Press.
- Robinson, John, Suzanne Bianchi, and Stanley Presser. 2001. *Family Interaction, Social Capital, and Trends in Time Use (FISCT), 1998-1999*: [Computer file] ICPSR version. College Park, MD: University of Maryland Survey research Center [producer], 1999. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
- Robinson, John. 1977. *How Americans Use Time: a Social Psychological Analysis of Everyday Behavior*. New York: Praeger.

Robinson, John. 1985. "The Validity and Reliability of Diaries Versus Alternative Time Use Measures. in *Time, Goods, and Well-Being*, F. Thomas Juster and Frank Stafford, eds. Ann Arbor, MI: Institute for Social Research, The University of Michigan.

Scheuch, Erwin K. 1972. "The Time-Budget Interview." In *The Use of Time*, ed. Alexander Szalai. The Hague: Mouton.

Table 1. Descriptive Statistics

		Mean	STD	Min	Max
Gaps (in minutes)					
Paid work	- excluding secondary activities	60.1	107.6	0.0	540.0
	--including secondary activities	63.8	115.8	0.0	600.0
	--only over-reported minutes, else 0	43.7	99.3	0.0	540.0
House work	- excluding secondary activities	126.2	122.7	0.0	420.0
	--including secondary activities	125.7	121.1	0.0	450.0
	--only over-reported minutes, else 0	112.1	123.8	0.0	420.0
TV viewing	- excluding secondary activities	94.8	118.2	0.0	620.0
	--including secondary activities	101.5	96.6	0.0	350.0
	--only over-reported minutes, else 0	50.9	99.3	0.0	620.0
Respondent traits					
	Years of education	13.4	2.3	7.0	18.0
	Income Categories	4.2	1.3	1.0	7.0
	Self-employed dummy	0.11	0.30	0.0	1.0
	Southern region dummy	0.33	0.45	0.0	1.0
	Age (or year of birth)	42.2	16.1	18.0	92.0
	Weekend dummy	0.27	0.43	0.0	1.0
	Number of children under 7	0.40	0.72	0.0	4.0
	Top 10 Metro area dummy	0.14	0.33	0.0	1.0
	Respondent feels rushed	0.31	0.44	0.0	1.0
Time use patterns					
Primary minutes	Paid work	215.5	246.8	0.0	875.0
	Housework	103.0	131.6	0.0	620.0
	TV viewing	119.9	158.6	0.0	910.0
Total intervals	Paid work	1.1	1.5	0.0	14.0
	Housework	2.3	2.3	0.0	17.0
	TV viewing	2.0	1.8	0.0	12.0
Secondary minutes	Paid work	7.2	50.2	0.0	600.0
	Housework	18.3	47.5	0.0	480.0
	TV viewing	49.1	96.2	0.0	795.0
Other	Number of precise times given	8.3	5.3	0.0	29.0
	Time interval standard deviation	116.9	48.4	42.4	354.0
Interviewer reports	Good memory of time diary	3.2	0.83	0.0	4.0
	Respondent gave begin/end times	0.60	0.47	0.0	1.0
	Reporting on diary or typical day	0.20	0.39	0.0	1.0

N=276 for all variables. Weighting is applied to all statistics.

Table 2. OLS Regression Predicting Paid Work Estimation Gap

Variable	Bivariate	Multivariate			
		Model 1	Model 2	Model 3	Model 4
Education	.04(.51)				
Income	-.16(.01)*	-.20(.00)*		-.22(.00)*	-.22(.00)*
Age	-.14(.02)*	-.15(.01)*		-.12(.03)*	-.13(.02)*
Self-employed	.18(.00)*	.21(.00)*		.20(.00)*	.20(.00)*
South	.05(.36)				
Number children under 7	.04(.54)				
Top 10 Metro Location	-.07(.25)				
Weekend	-.13(.03)*	-.14(.02)*		-.06(.32)	-.06(.33)
Primary paid time	.25(.00)*		.14(.07)+	.22(.00)*	.16(.01)*
Secondary paid time	.37(.00)*		.32(.00)*		.34(.00)*
Total paid intervals	.26(.00)*		.10(.22)		
Precise times given	-.05(.44)		-.09(.14)		
Time interval STD	-.03(.59)		-.07(.28)		
Good memory	.01(.82)				
R gave begin/end times	.05(.40)				
R gave typical day est.	.05(.38)				
House work Gap	.04(.54)				
TV Gap	.07(.25)				
R-squared		0.11	0.19	0.15	0.26

N=276 for all analyses. Estimates are standardized. P-values are in parentheses. *p<.05 +p<.10

Table 3. OLS Regression Predicting Housework Estimation Gap

Variable	Bivariate	Multivariate			
		Model 1	Model 2	Model 3	Model 4
Education	-.25(.00)*	-.22(.00)*		-.15(.01)*	-.14(.02)*
Income	-.17(.01)*	-.09(.13)		-.10(.08)+	-.10(.09)+
Age	-.10(.11)				
Self-employed	.04(.53)				
South	.02(.73)				
Number children under 7	.27(.00)*	.26(.00)*		.27(.00)*	.27(.01)*
Top 10 Metro Location	.11(.07)+	.05(.34)		.05(.39)	.06(.31)
Weekend	.08(.17)				
Primary housework time	.26(.00)*		.32(.00)*	.21(.00)*	.21(.00)*
Secondary housework time	.17(.00)*		.16(.01)*	.11(.04)*	.11(.05)*
Total housework intervals	.11(.08)+		-.11(.25)		
Precise times given	-.12(.06)+		-.02(.78)		
Time interval STD	.11(.06)+		.14(.03)*	.21(.00)*	.20(.00)*
Good memory	-.14(.01)*				-.05(.42)
R gave begin/end times	-.12(.04)*				-.00(.96)
R gave typical day est.	-.07(.26)				
Paid work gap	.04(.54)				
TV Gap	.28(.00)*				
R-squared		0.15	0.11	0.23	0.23

N=276 for all analyses. Estimates are standardized. P-values are in parentheses. *p<=.05 +p<.10

Table 4. OLS Regression Predicting Television Viewing Estimation Gap

Variable	Bivariate	Multivariate			
		Model 1	Model 2	Model 3	Model 4
Education	-.27(.00)*	-.27(.00)*		-.10(.02)*	-.10(.03)*
Income	-.11(.06)+	-.02(-.31)		.01(.65)	.02(.65)
Age	.01(.82)				
Self-employed	-.03(.61)				
South	.09(.15)				
Number children under 7	.16(.01)*	.16(.01)*		.24(.00)*	.24(.00)*
Top 10 Metro Location	.06(.34)				
Weekend	-.00(.97)				
Primary television time	.57(.00)*		.64(.00)*	.62(.00)*	.62(.00)*
Secondary television time	.43(.00)*		.47(.00)*	.45(.00)*	.45(.00)*
Total television intervals	.47(.00)*		-.02(.72)		
Precise times given	-.15(.02)*		.01(.86)		
Time interval STD	.19(.00)*		-.10(.06)+	-.06(.17)	-.06(.17)
Good memory	-.11(.07)+				-.00(.91)
R gave begin/end times	-.06(.33)				
R gave typical day est.	.06(.36)				
Paid work gap	.07(.25)				
Housework gap	.28(.00)*				
R-squared		0.10	0.53	0.60	0.60

N=276 for all analyses. Estimates are standardized. P-values are in parentheses. *p<=.05 +p<.10

Appendix A. Diary Questions

Next, I would like to ask you about the things you did **yesterday**. I want to know **only** the specific things you did yesterday, **not** the things you usually do. Let's start at midnight [fill day of week before diary day], that is, the night **before** last.

Q1) What were you doing [fill day of week before diary day] at midnight?

RECORD ACTIVITY CODE

0 OTHER (must specify

activity): _____

1 BATHING/SHOWERING 4 SHOPPING 7 TRAVELING

2 DRESSING/GROOMING 5 PREPARING MEALS/SNACKS 8 WATCHING TV

3 EATING MEALS/SNACKS 6 SLEEPING/NAPPING 9 WORKING FOR PAY

** IF PERSON REPORTED TRAVELING ASK QUESTION Q2B

Q2A) Where were you? _____

0 OTHER (must specify

where): _____

1 HOME 5 GROCERY STORE

2 OTHER'S HOME 6 OTHER STORE/MALL

3 OUTDOOR AWAY FROM HOME 7 SCHOOL

4 OFFICE BUILDING/FACTORY 8 RESTAURANT/BAR

Q2B) How were you traveling? _____

1 CAR/TRUCK/MOTORCYCLE 2 BUS/TRAIN/AIRPLANE 3 WALKING 0 OTHER

Q3) What time did you finish? _____ _m

_____ SECONDARY

ACTIVITY _____

IF RESPONDENT IS SLEEPING, BATHING OR SHOWERING GO TO NEXT ACTIVITY

Q4) At any time while you were (REPEAT ACTIVITY) did you do anything else? (like

talking, reading, watching tv, listening to the radio, eating or caring for children)

RECORD ACTIVITY CODE: _____ **IF MORE THAN ONE, REPORT AS OTHER

0 OTHER (must specify

activity(ies): _____

1 NO/NOTHING ELSE 5 EATING MEALS/SNACKS

2 WATCHING TV 6 CHILD CARE

3 TALK TO OTHERS 7 READING

4 LISTEN TO MUSIC OR RADIO

Q5) While you were (REPEAT ACTIVITY) who was with you: _____

0 OTHER (must

specify): _____

1 ALONE/NO ONE 5 CO-WORKERS

2 SPOUSE ONLY 6 FRIENDS

3 CHILD(REN) ONLY 7 RELATIVES

4 SPOUSE & CHILD(REN) 8 STRANGERS/CROWD

Q6) What did you do next? _____

0 OTHER (must specify

activity): _____

1 BATHING/SHOWERING 4 SHOPPING 7

TRAVELING

2 DRESSING/GROOMING 5 PREPARING MEALS/SNACKS 8

WATCHING TV

3 EATING MEALS/SNACKS 6 SLEEPING/NAPPING 9 WORKING

FOR PAY

After Q6, same logic is repeated until all activities are obtained.