CrossMark

ORIGINAL PAPER

Long Working Hours and Well-being: What We Know, What We Do Not Know, and What We Need to Know

Daniel C. Ganster[1] · Christopher C. Rosen[2] · Gwenith G. Fisher[3]

Abstract Despite a widespread belief in both the academic and public policy literatures that working long hours is deleterious to health and well-being, our critical review of this large and complex literature fails to support a robust direct causal effect of work hours on either physical or mental well-being outcomes. Large-scale epidemiological studies, many of which are prospective and include objective health outcome measures, support a statistically significant association between long work hours and coronary heart disease and depression, but the effect sizes are very small. Moreover, there is an absence of true longitudinal studies that assess the consistency of working long hours over time and its relationship to well-being. Our review suggests that the effects of working long hours are nuanced in that they may vary considerably for different working populations based on gender, age, working conditions, and other factors. Primary and meta-analytic studies suggest that such moderator effects are plausible, yet rigorous testing of these remains to be done. We conclude with suggestions for specific moderator effects that seem worth investigating in future research.

✉ Christopher C. Rosen
   crosen@walton.uark.edu

[1] Department of Management, Colorado State University, Fort Collins, CO, USA

[2] Department of Management, Sam M. Walton College of Business, University of Arkansas, Fayetteville, AR, USA

[3] Department of Psychology, Colorado State University, Fort Collins, CO, USA

Work hours are an aspect of working conditions that may vary both within, as well as between, jobs and that has the potential to impact worker well-being. More explicitly, it has long been assumed that excessive work hours impose a high cost to workers in terms of overall quality of life, including physical and mental health (Barnett 2006). The presumed consequences of long work hours are so important that reducing them and establishing clear policies about work hours was among the first goals for which organized labor battled in the USA. Moreover, work hours is one of the few aspects of working conditions that has been regulated by national governments and labor unions. Recently, for example, this issue received attention when Sweden reduced the work day to 6 h in an effort to increase work/life balance and make workers happier (BBC 2016). Likewise, professional associations have also begun to establish guidelines aimed at limiting the number of hours worked by groups that they represent. For example, the Accreditation Council for Graduate Medical Education (ACGME) established guidelines aimed at limiting the weekly work hours of medical residents (Landrigan et al. 2008).

Acknowledging the potential hazards to the health of workers as well as risks posed to customers and the public from overly fatigued workers, the National Institute for Occupational Safety and Health (NIOSH) convened a National Occupational Research Agenda (NORA) Long Hours Work Team to articulate research needs regarding long working hours and their effects on health and safety (Caruso et al. 2006). Since the publication of this agenda, numerous large-scale studies have considered how working longer work hours than the typical 35–40-h workweek relates to a variety of mental and physical health outcomes. Segments of this diverse literature, ranging from psychology and management to epidemiology, have previously been summarized in both qualitative (e.g., Bannai and Tamakoshi 2014; Solovieva

et al. 2013) and quantitative reviews (Kang et al. 2012; Kivimaki et al. 2015a, b; Virtanen et al. 2012a, b).

In the current paper, we aim to critically assess this growing body of literature, with a particular focus on (a) meaningful effect sizes, (b) confounding factors in the extant research, and (c) potential moderators and mediators of the effects of long work hours on worker well-being. Our interest is on the well-being of employees who work long hours and not on that of their coworkers or customers and clients. Our primary goal is to suggest a set of key research questions and methodological approaches that address what we identify as important gaps in the extensive work hours literature.

Working Hours

The study of working hours would seem to be a straightforward enterprise. Compared with many working conditions and work-related stressors that have dominated the occupational health and organizational literature—such as role overload, lack of control, injustice, to name a few (Ganster and Rosen 2013)—the construct of working hours is a conceptually simple one. It refers to an objective variable that can potentially be measured directly. When respondents are asked to report their own work hours, errors may be mostly ones of recall or from certain biases such as a desire to hide income from the IRS, social desirability, or other motivations (Robinson et al. 2002). For more conceptually complex constructs, these sources of error also exist and are compounded by a host of other measurement issues, not the least of which is the fact that respondents must engage in a subjective appraisal process to report their experience. For example, self-reports of work overload, a commonly studied construct in the work stress arena, reflect objective role demands—both qualitative and quantitative—as well as the respondent's ability to handle those demands, resulting in a relatively low level of convergence between self and observer reports of workload (Ganster 2005; Spector et al. 1988). In addition, the concept of work hours, as a putative cause of well-being, is easily understood by employers, workers and their advocates, as well as policy makers. If long work hours are proven to threaten mental and physical well-being, intervention strategies and policy recommendations are generally clear, and their cost/benefit ratios can be estimated (Goh et al. 2016).

The most common underlying causal assumption in the work hours literature is that long work hours exert a negative effect on well-being. There are several reasons why long work hours might impair well-being. First, hours spent working are subtracted from time available for non-work activities (e.g., family obligations and leisure pursuits), creating conflicts between the two domains. For example, in their seminal article on work/family conflict, Greenhaus and Beutell (1985) identified time-based conflict as a key source of work/family

conflict. As Barnett (2006, p. 104) notes, "both workplaces and families are greedy institutions.". Such conflicts can trigger a cascade of social consequences, from short-term stress to longer-term deterioration of social relationships, which in turn can affect health and happiness. Second, long working hours induce physical wear and tear and lessen the opportunity for recovery from fatigue (Fritz and Sonnentag 2006). Third, excessively long work hours might lead to poorer diet and exercise patterns, and increased smoking and alcohol consumption, thereby providing another potential pathway to poor health (Solovieva et al. 2013; Steptoe et al. 1998). Fourth, fatigue induced by long work shifts, or insufficient recovery from work shifts, can lead to more accidents causing personal injuries that eventually lead to chronic health outcomes. Fifth, fatigue or sleep deprivation from long work shifts may lead to the depletion of personal resources that are necessary for executive functioning, thus having a negative impact on job performance, safety, and other important organizational outcomes (Christian and Ellis 2011; Landrigan et al. 2008; West et al. 2009). Finally, long hours also prolong the individual's exposure to whatever psychosocial or physical stressors might be on the job, and thus increase the harmful effects of other negative working conditions.

One must also consider the counter hypothesis that longer work hours, when they lead to higher incomes, status, and/or greater social support, might also produce positive outcomes. Such outcomes might partially accrue from increased benefits, job security, and opportunities for advancement as well (Barnett 2006). As Barnett (2006) noted, however, little attention has been paid to this hypothesis or to how the benefits of long work hours might counterbalance the negative effects of long hours.

Review of Meta-analyses of Work Hours and Well-being

A number of meta-analyses have summarized research on the relationship between work hours and various measures of well-being, which we broadly define to include physical health and psychological well-being. Outcomes considered in these meta-analyses include indicators of coronary heart disease (CHD), diabetes, obesity, mental health, and health-related behaviors. Because multiple meta-analyses have already been reported, we chose to review and compare the existing ones rather than create yet another one.

Long Work Hours and CHD One of the first meta-analyses to include CHD outcomes of long work hours was reported by Sparks et al. (1997) almost 20 years ago. Sparks et al. (1997) included studies that reported weekly work hours and excluded those that focused on extended work shifts (e.g., 12-h shifts). This selection resulted in an analysis of 19 studies and 21 independent samples comprising 37,623 individuals. The sample sizes in each study ranged from 32 to 7382, and

the publication dates of the studies ranged from 1958 to 1997. Sparks et al. (1997) reported significant, albeit small, relationships between work hours and health, with the average correlation for overall health (all indicators combined) being .13. When they separated health outcomes into physiological (which included some CHD outcomes) and psychological indicators, they found a higher weighted average correlation for psychological outcomes ($r = .15$) than for physiological ones ($r = .06$). In both cases, sampling error variance accounted for relatively little of the variability across studies (7.39% for psychological and 21.89% for physiological), suggesting the presence of one or more moderator variables that may explain the relation between work hours and health outcomes. Because of the small number of studies, Sparks et al. (1997) did not attempt moderator analyses. The studies in Sparks et al.'s analysis represented a broad array of outcome variables, moreover, ranging from reports of coronary heart disease to irritability/tension, problems with relationships, and symptoms of carpal tunnel syndrome. As such, the broad array of variables examined in the study may serve as at least one plausible explanation for the range of effect sizes.

More recently, meta-analyses have focused on specific health outcomes. Three such analyses focused exclusively on coronary heart disease and/or stroke. The earliest of these (i.e., Virtanen et al. 2012a, b) included 12 studies that investigated an outcome classified as cardiovascular disease, including stroke, but excluded studies that focused on risk factors such as hypertension, metabolic syndrome, or diabetes. The total sample size of the studies was 22,518. Results were presented as relative risk (RR) factors, which indicate the ratio of the probability of having a particular health condition to the probability of not having the condition. In the combined sample, the minimally adjusted RR, adjusted for age and sex, was 1.80 (95% CI, 1.42, 2.29), whereas the relative risk average from those studies with more control variables (e.g., socioeconomic position) was 1.59. (95% CI, 1.23, 20.7). Virtanen et al. (2012a, b) also performed subgroup analyses, separating studies into either prospective cohort or case-control designs. Although only one of the four prospective cohort studies reported a significant relationship between work hours and CHD, the overall weighted relative risk was 1.39 (95% CI, 1.12, 1.72). The result for the seven case-control studies was higher at 2.43 (95% CI, 1.81, 3.26). As the authors noted, however, such studies rely on retrospective reports of working hours, which can produce recall bias and reverse causality. Overall, the most conservative conclusion regarding the prospective studies suggests a significant, albeit small, effect of long working hours, with those working long hours having about a 39% greater chance of subsequently suffering some form of coronary heart disease than those working "normal" hours.

Kang et al. (2012) also examined studies that assessed the relationship between long working hours and cardiovascular disease. Kang et al.'s (2012) meta-analysis included 11 studies with a combined sample size of 15,923 (12,561 from prospective cohort studies and 3362 from case-control studies). Of the 11 studies they included, 6 were in common with the meta-analysis reported by Virtanen et al. (2012a, b). The combined relative risk of long working hours was 1.37 (95% CI, 1.11, 1.70), and the relative risk in the case-control studies (1.43) and the prospective cohort studies (1.41) were similar. Overall, then, Kang et al.'s (2012) conclusions are essentially the same as those from Virtanen et al. (2012a, b).

Kivimaki et al. (2015a) published by far the largest meta-analysis of the association between working hours and cardiovascular disease. Kivimaki et al. reviewed 25 studies, 20 of which were unpublished, with a combined sample size of 603,838 individuals who were free from CHD at baseline and who were followed up for a mean of 8.5 years. For stroke as a separate outcome, they reviewed 17 studies with a combined sample size of 528,908 individuals who were free from disease at baseline and were followed up for a mean of 7.2 years. The average risk ratio for CHD for those who worked more than 55 h/week, compared with those working 35–40 h/week was 1.13 (95% CI, 1.02, 1.26). This result reflects adjustment for age, sex, and socioeconomic status. The average risk ratio for stroke was somewhat stronger at 1.33 (95% CI: 1.11, 1.61). Both of these were significant, but CHD only at $p = .02$. Moreover, this analysis did not only omitted several studies that were included in earlier reviews by Kang et al. (2012) and Virtanen et al. (2012a, b) but also included some large studies published since those reviews. Most significantly, the authors obtained access to 20 large datasets which they were able to analyze at the individual level and then combine these analyses into their meta-analysis. Because they had access to so much individual-level data, they were able to control for additional, potentially confounding variables, including body-mass index, physical activity, and alcohol consumption.

Kivimaki et al. (2015a) meta-analysis has several unique strengths. First, they considered only prospective cohort studies in which individuals did not have CHD at the baseline measurement. Eliminating participants with CHD at baseline rules out reverse causality effects (e.g., that sick people would work fewer hours), as well as retrospective biases possible in case-control studies. Second, despite their exclusion of non-cohort designs, their combined sample size was many times larger than those of prior meta-analyses. Finally, because they had access to many datasets which they could analyze themselves, these researchers were able to control for a consistent set of potential confounders across studies.

That said, how does their effect size of working hours compare with the earlier meta-analysis of Sparks et al. (1997), which focused on weighted average correlations as measures of effect sizes? This comparison is complicated by the fact that in the meta-analyses of incident CHD or stroke, effect sizes

are expressed in terms of relative risk. Thus, those working longer hours (i.e., greater than 55/week) have a 13% greater risk of developing CHD over the next 8.5 years and have a 33% greater risk of experiencing a stroke over the next 7.2 years.

How do these relative risks compare with the small, but significant, average weighted correlation of .06 from Sparks et al. (1997), which was statistically significant at $p < .05$? Bonett (2007) reviewed several approaches for calculating approximate correlations from odds ratios. Using the tetrachoric correlation approximation of Digby (1983), the meta-analytic result from Kivimaki et al. (2015a) yielded an approximate tetrachoric r of .05. As with the small average r reported by Sparks et al. (1997), this very small effect size derived its statistical significance from the large combined sample size of the meta-analysis. Thus, at least in terms of CHD, the most robust meta-analytic findings compare favorably with earlier ones reported for a much broader category of physiological outcomes and indicate that there is a significant, but small, association between long work hours and CHD.

Another way of looking at this issue is to consider the incident rates (the basis for RR calculations) from Kivimaki et al. (2015a). The overall incident rate over 8.5 years was .8% (4768 events/603,838 individuals). Multiplying this rate by the RR of 1.18 yields an approximate rate for those working long hours of .9%.[1] The effect size for stroke incidence is somewhat higher, with an approximate tetrachoric r of .11, but still small in practical terms. Translating the RR to an incident rate for those working long hours, the risk of developing stroke in 7.2 years was .3% for the entire sample, and .4% for the high work hours group.

Long Work Hours and Diabetes and Obesity In addition to CHD, several studies have examined the association of long work hours to other physical health conditions that can also be considered risk factors for CHD (e.g., obesity and hypertension), as well as serious diseases in their own right (e.g., diabetes). The US Centers for Disease Control reported that diabetes (95% of which is type 2) was the seventh leading cause of death in the USA in 2013, and that more than 20% of health care spending is for people diagnosed with diabetes (https://www.cdc.gov/chronicdisease/resources/publications/aag/diabetes.htm). Kivimaki et al. (2015b) reported a large-scale meta-analysis of prospective cohort studies that examined the risk of developing type 2 diabetes from working long hours. Their analysis combined the meta-analysis of 4 published studies with the results from 19 unpublished datasets representing 222,120 individuals with an average follow-up period of 7.65 years. As with Kivimaki

et al. (2015a), Kivimaki et al. (2015b) were able to control for a fairly broad list of confounds with the unpublished datasets, including age, sex, SES, body-mass index, physical activity, and alcohol consumption. The overall effect for the combined sample was not significant (RR = 1.07; 95% CI, .89, 1.27). However, in a subgroup analysis separating samples into low and high socioeconomic status (SES) groups, they found a significant association (hypothesized) for those in the low SES group (RR = 1.29; 95% CI, 1.06, 1.57). Whereas there was significant heterogeneity in results for the entire sample, there was no longer significant heterogeneity in the subgroups. The tetrachoric r approximation for the low SES group is .10. Kivimaki et al. (2015b) expressed this effect size as 13 additional cases of diabetes in 10,000 person years for low SES workers compared with the high SES workers.

The review by Solovieva et al. (2013) examined studies testing the association between psychosocial working conditions and long work hours and various obesity or weight gain outcomes. Although this was not an actual meta-analysis, Solovieva et al. (2013) reported that across 14 studies about 40% of the tested relationships between work hours and a weight-related outcome were statistically significant. Overall, the studies showed an inconsistent pattern of weak associations, with the strongest being for weight gain associated with overtime work for men. Unfortunately, the authors provided no effect size estimates.

Finally, a recent study by Brummelhuis et al. (2016) shows that 10 of 763 employees found that long work hours was not predictive of risk factors of metabolic syndrome (high blood pressure and elevated cholesterol levels). The only significant predictor of metabolic syndrome was a workaholism measure, and even that was related to metabolic syndrome only for workers whose engagement was low.

Summary of Long Work Hours and Physical Health To summarize, there is a large literature that addresses the relationship of long work hours and a host of physical health conditions, especially coronary heart disease. Moreover, this literature has many prospective cohort designs with long follow-up times, some of which control for an extensive list of potentially confounding variables. Most authors of these meta-analyses conclude that working long hours poses a significant risk for CHD. We argue, however, that very large sample sizes, especially aggregated samples from meta-analyses, possess a high level of power to detect even very small effect sizes that might not be practically meaningful. Our assessment of the effect sizes from the large epidemiological studies suggests that they are indeed quite small and comparable with the correlational effect sizes reported by Sparks et al. (1997) in their early meta-analysis. One other observation we would make about this literature is that most meta-analyses report significant levels of between-study variance in effect sizes over and above that accounted for by

---

[1] Actually, this somewhat overestimates the impact because the RR figure should be applied to the incident rate of the normal work hours group, which was unavailable, rather than that of the entire sample.

sampling error. This is a point we return to later when we suggest a research agenda for addressing long working hours. The following points summarize our conclusions regarding work hours and CHD:

- There is a significant, but small, association between long work hours and CHD.
- There is weak and inconsistent evidence supporting a relationship between long work hours and diabetes/obesity.
- Meta-analyses indicate significant between-study variance in effect sizes across studies examining the relationship between work hours and physiological health outcomes.

Long Work Hours and Mental Health and Health-Related Behavior In this category, we include a broad array of psychological and behavioral health outcomes ranging from subjective distress indices to depression and alcohol consumption. The Sparks et al. (1997) meta-analysis represents just such a broad sampling of psychological outcomes, with dependent variables such as subjective stress, anxiety, depression, irritability, hostility, somatization, lack of concentration, and social dissatisfaction. Aggregating these outcomes resulted in a weighted average effect size ($N = 35,445$) of .15. Considering the self-report nature of most of the outcome measures coupled with self-reported work hours, these effect sizes might be somewhat inflated (due to common method variance) relative to those that studied CHD outcomes (Johnson et al. 2011). Sparks et al. (1997) also calculated that most of the variance between study results (93%) was not accounted for by sampling error. Because they concluded that there were too few studies available to conduct moderator analyses, they did not attempt any.

Unlike the literature on physical health outcomes, studies on psychological outcomes have not been considered in as many meta-analyses. One meta-analysis, however, was reported by Ng and Feldman (2008), who surveyed a wide range of studies that examined both putative antecedents and outcomes of long work hours. They meta-analyzed 38 studies ($N = 21,280$) that examined self-rated "mental strain," which consisted of low scores on self-rated mental health. The weighted average correlation of these studies was .06 ($p < .05$). The 23 studies that measured "job stress" ($N = 16,268$) had a somewhat larger average correlation of .13, although job stress is generally not considered a mental health outcome, per se.

More recently Bannai and Tamakoshi (2014) conducted a qualitative review of epidemiological studies that reported mental health outcomes, which included depression (5 studies) and anxiety (2 studies). Their conclusion was that working long hours was associated with both depression and anxiety. Because theirs was a qualitative review, we review

these individual studies in order to glean some effect size estimates.

Nagashima et al. (2007) reported the results of a cross-sectional survey of 715 male day workers in a Japanese chemical factory. Participants self-reported working hours as well as depression and anxiety. Nagashima et al. (2007) divided the sample into work hour categories and computed odds ratios with dichotomous outcomes using arbitrary cutoff points of the continuous scales. Of the six working hours groups (ranging from less 200 h/month to over 280 h/month), only one (260–279 h/month) had a significant odds ratio (2.75). This category was the second highest in terms of working hours. For anxiety, the two highest categories of working hours showed significant odds ratios (2.28 and 2.51). The researchers had controlled for age, marital status, smoking habit, alcohol consumption, and exercise habit.

Virtanen et al. (2012b) conducted a prospective cohort analysis with 2123 full time workers from the Whitehall II study of British civil servants. The sample was free of psychological morbidity at baseline and participants were classified 7.2 years later on the basis of whether they experienced a major depressive episode or not, based on a diagnostic interview scored according to DSM-III-R criteria. The Virtanen et al. (2012b) analyses were conducted on 4 working hours groups (7–8, 9, 10, 11–12 h/day, respectively) using a logistic regression for the dichotomous depression outcome. There was no significant effect for working hours until controlling for SES, at which point the odds ratio for those working 11–12 h days was 2.43 compared with those working 7–8 h days. This association held up after also controlling for physical disease at baseline, smoking, alcohol use, job strain, and social support. SES seemed to exert a suppressor effect, likely because the relation between work hours, SES, and depression is complex: high SES individuals tend to work longer hours, and SES is negatively associated with depression. Overall, given the cohort design, the extensive controls, and the relatively objective assessment of major depressive episode, this is a strong study. As for effect size, this might be best illustrated by comparing the incidence rates for the low and high work hour groups. The rate for the 7–8-h group was .034, whereas the rate for the 11–12-h group was 4.4. Given the large sample size, the odds ratio (with controls) was statistically significant at $p < .05$. In comparison, the median-reported population 1-year depression prevalence rate is often near 5% (Eaton et al. (2008). Virtanen et al. (2012b) speculated that their relatively low rates might be due to their including only participants who did not have psychological problems at baseline.

Virtanen et al. (2011) also examined a sample from the Whitehall II study. In this case, the baseline year was 1997–1999 (phase 5), which was the follow-up phase for the Virtanen et al. (2012b) study discussed previously. In the current study, the researchers followed up individuals who were free from depression ($N = 2549$) and anxiety symptoms

(N = 2618) at baseline and assessed them again for depression and anxiety a few years later at 2001 and at 2002–2004 (phases 6 and 7). Work hours were assessed once at baseline. Their results compared groups working 35–40, 41–55, and over 55 h/week. Overall, there was a significant difference between the over 55 group compared with the 35–40 group (hazard ratio = 2.14), but a significant interaction effect with sex showed that this effect held only for women. As with Virtanen et al. (2012b), this study controlled for several factors at baseline. The largest effect size (women with all confounders controlled for) is reflected in the incidence rates, which were 5.3/100 person years for the over 55-h group compared with 2.2/100 person years for the 35–40-h group.

It is difficult to integrate these two studies from the Whitehall II project, as they both examined individuals from the same study. In the second study (i.e., Virtanen et al. 2012b), the authors used a different assessment for depression (the University of Michigan Composite International Diagnostic Interview (UM-CIDI)) at phase 5 to determine depression "caseness," whereas in the first study items from the General Health Questionnaire (GHQ-30) Depression scale were used as the baseline depression measure at phase 5 and to determine caseness at later phases. In the second study, the researchers used an earlier measure of work hours from phase 3 that the authors described as "not as accurate as at phase 5 including hours worked only on an average weekday" (Virtanen et al. 2011, p. 2487). It is also not apparent why the second study by Virtanen et al. (2012b) did not analyze sex interaction effects when these were so prominent in the earlier study, except that they might have had too little power to do so because of the small number of depression events identified by the UM-CIDI relative to the GHQ. Virtanen et al. (2012b, p. e30719) noted that since the GHQ "also detects a range of minor psychiatric disorders, such as subclinical depression, it is possible that our baseline exclusion of GHQ-30 cases is over zealous." The GHQ-30 served as the outcome measure for the first study by Virtanen et al. (2011) analyses.

A prospective cohort study by Shields (1999) followed up 3783 Canadian respondents from the National Population Health Survey who worked at least 35 h/week in 1994/1995 (baseline year) and who provided data in the follow-up year (1996/1997). Work hours were self-reported for the year preceding the baseline. Experience of a major depressive disorder in the year preceding the follow-up survey was determined by questions from the Composite International Diagnostic Interview (World Health Organization 1994) that were scored and transformed into a probability estimate of a diagnosis of major depressive disorder. Odds ratios were calculated for those working more than 41 h/week relative to those working 35–40 h. Odds ratios were adjusted for age, irregular work shift, white collar job, marital status, children, education, household income, and work stress and were computed

separately by sex. Women working long hours had a significantly higher rate of depression than those working regular hours (odds ratio = 2.2; 95% CI, 1.1, 4.4), but this was not the case for men. Despite the seemingly small difference in working hours based on the weekly cutoff, men in the high hours group averaged 55 h/week, whereas women in this category averaged 51 h/week.

Driesen et al. (2010) analyzed the prevalence of depression, measured with a single screening item, for 4951 men and 1680 women from the Maastricht Cohort Study. Using those who worked on the day shift, they compared depression rates for those working 36–40 h, more than 40 h, 26–35 h, and less than 26 h/week. Men working more than 40 h per week (i.e., the high work hours group) had a significantly *lower* rate of depression than the referent 36–40 h/week group (odds ratio = .74; 95% CI, .56,.99), while those who worked fewer than 26 h/week had a significantly *higher* rate of depression (odds ratio = 2.90; 95% CI, 1.37, 6.15). Odds ratios were computed after adjusting for a long list of potential confounding variables, including social support, job demands, and control. There were no significant effects for women. Driesen et al. (2010) conducted a separate analysis of overtime work on 4618 men and 690 women. Comparing groups who worked (a) fewer than 5 h/per week overtime, (b) 5–10 h, (c) 11–20 h, and (d) more than 20 h, they found an increased risk for depression in men for those in the 11–20 and over 20 h/week groups compared with the fewer than 5 h/week group, but the statistical significance of these differences was erased when the full set of control variables was added. There were no significant overtime effects for women.

Another cross-sectional study of working hours and depression that was not covered in the Bannai and Tamakoshi (2014) review was reported by Amagasa and Nakayama (2012) who obtained survey data from two businesses in Japan (N = 218 and N = 1160). In addition to work hours per week and depression, they also collected measures of job demands and control. Analyzing the data separately for each company, they first fitted SEM models that used different configurations of the variables to predict depression scores. They then took a "best-fit" model and calculated odds ratios on different groups implied by the model. The smaller sample yielded no significant results from the logistic regression. The best-fit model for the larger sample specified working hours as an exogenous variable that led to job demands, which, in turn, led to depression. Control was also specified as an exogenous variable with a direct path only to depression. Despite there being no interaction effects specified in their SEM model, the resulting logistic regression reflected groups of low and high job demands within low (less than 60 h/week) and high (more than 60 h/week) work hour categories. Two significant odds ratios involving work hours resulted from this analysis. One indicated an increased risk for those in high-demand jobs working low hours (OR = 1.60; 95% CI, 1.20, 214), and the

other indicated an increased risk for those working in high-demand jobs and working long hours (OR = 2.24; 95% CI, 1.02, 4.90). These results imply a main effect for job demands (one was also reported for control), but it is difficult to derive a real effect for work hours, per se. We thus consider the results from this study to be ambiguous regarding work hours and risk of depression.

Summary of Work Hours and Mental Health The two meta-analyses covering studies that investigated the relationship of long work hours to mental health (Ng and Feldman 2008; Sparks et al. 1997) reported very small effect sizes for broad categories of mental health indicators. Later, large-scale epidemiological studies focused more narrowly on specific mental health outcomes such as depression and anxiety. Some of these were cross-sectional studies that reported significant associations for depression. The more rigorous prospective cohort studies showed a much less consistent pattern. Two analyses from the Whitehall II studies (Virtanen et al. 2011, 2012b) reported higher risk ratios for working longer than normal work hours, which are either quite small or hold only for women. Similarly, the prospective study by Shields (1999) found a significant effect, but again, only for women. The prospective cohort study by Driesen et al. (2010) actually found a *negative* relationship between long work hours and depression for men and no association for women. Thus, when considering the strongest studies, it is difficult to make a strong case that working long hours will increase the rates of depression or anxiety. The following summarizes our assessment of the work hours and psychological well-being literature:

- Evidence supporting a relationship between long work hours and psychological well-being is equivocal.

Methodological Issues

The accumulated evidence, especially the effect sizes produced by large prospective studies, suggests that *long working hours, by themselves, have a very small impact on both physical and mental health outcomes*. This conclusion is more circumspect than that advanced by many of the researchers and reviewers in this literature. For example, Goh et al. (2016) recently estimated, after a long and complicated analysis, that the marginal costs of long work hours were about 13 billion dollars in the USA. This estimate was derived after taking into account a number of other contributors to healthcare costs, ranging from low support ($9 billion) to high job demands ($48 billion). In their view, long work hours and shift work (estimated separately) "have profound health consequences ..." (Goh et al. 2016, p. 610). Many other reviewers of this literature conclude that long work hours pose a

significantly increased risk for CHD, stroke, and depression. In this section, we discuss key methodological issues that might be obscuring the true effect sizes of long working hours.

Measurement of Work Hours Compared with many other potential work stressors, working hours is a relatively straightforward construct. By far the most common method for measuring working hours is self-report. As we noted above, self-report errors mostly concern recall issues and some bias for both under- and/or over-reporting. But these self-report limitations are no more, and we would argue are less, severe than for many other job demands (e.g., work overload, low control, abusive supervision, lack of support, and injustice). The fact that many of the best prospective cohort studies assess working hours at baseline with a healthy cohort and then relate them to oftentimes objective health outcomes years later serves to mitigate many of the self-reporting bias effects. Of course, this also applies to other self-reported stressors, which tend to have much more significant associations with both mental and physical health outcomes (Goh et al. 2016).

There are complications with the measurement of long work hours, however. For some workers, measures of work hours may be criterion deficient because the measure does not fully account for all hours worked. For example, ideally one would want the measure to reflect the total number of working hours across all jobs for those who hold multiple jobs, and to include time spent working at home as well. Given technological advances that have made it easier for individuals to work anywhere at any time and blurring boundaries between work and home (Kossek et al. 2012), respondents may not necessarily report all hours spent working. For the most part, however, investigators attempt such a summative measure, so it is unlikely that an under-reporting of work hours is systematically biasing associations downward.

The large prospective cohort studies, despite generally having continuous measures of working hours, invariably create multiple categories of high and low working hour groups. Sometimes, there are as many as five or six categories, but often the data are dichotomized into high and low groups. Moreover, the arbitrary cutoff values tend to vary across studies. Some dichotomize at 40 h/week (e.g., Shields 1999), some at 45 h (e.g., Tarumi et al. 2003), and one as high as 55 h/week (Amagasa and Nakayama 2012, cutoff based on 240 h/month). Whatever the cutoff chosen, investigators discard variance and lose precision when they artificially convert continuous measures into nominal categories.

Perhaps the biggest measurement limitation in the literature on long working hours, however, derives from the lack of any true longitudinal studies—that is, those that assess working hours at multiple time points. Despite the relatively large number of prospective cohort studies, we could find none that measured work hours other than at baseline. The study by Shields (1999) might be considered an exception, had she

utilized the data she had available for the prediction of depression. As we noted above, Shields' (1999) study followed up a prospective cohort design over a 2-year period, but unlike other such studies, she assessed work hours at both time periods. She reported analyses of depression, however, that relied solely on the baseline level of work hours, despite finding some effects for the *change* in work hours from baseline to follow-up on health risk factors such as weight gain and smoking. The Whitehall II study of British civil servants is another large-scale panel study spanning many years that offers an excellent opportunity to longitudinally investigate the effects of change and/or consistency in working hours on health outcomes. Virtanen et al. (2011, 2012b), reviewed above, are two analyses from this data that each focused on certain time periods and did baseline-only analyses of working hours, even though they were measured at multiple time points.

What is the importance of measuring working hours at more than just baseline? First, self-reports of working hours typically ask respondents to report the hours worked "in a typical workweek." The typical workweek, however, could certainly change over time periods constituting multiple years. Shields (1999), for example, found that although many of her respondents reported the same level of working hours over 2 years (64% of men and 69% of women who worked standard hours at baseline continued to so at follow-up, and 66% of men and 48% of women who worked long hours at baseline continued to do so at follow-up), there was certainly a significant amount of variability over time, especially for women. Second, using measures of work hours for multiple time periods affords the investigator opportunities to examine the effects of change in addition to the effects of consistent high or low levels of work hours. An example of how consistency of exposure to other working conditions can really matter is evident in a study by Bosma et al. (1997). Also using data from the Whitehall II study, Bosma et al. measured job control at baseline and at approximately 3.7 years later and used these two measures to predict cardiovascular disease about 4 years after that. What they found was that those who had stable and high levels of control over the years showed the lowest level of risk compared with those whose control was consistently low or had changed one way or the other over time. Thus, available evidence suggests that (a) there is significant change in work hour patterns in as short as a year, and (b) at least with respect to other job conditions, consistent high levels of exposure seem to matter.

Finally, it is important to consider the duration of exposure to long working hours. It is rare for any study to assess the effects of long work hour exposures of more than just a few years. It may be unlikely that such relatively short exposure periods are sufficient to lead to CHD. For this reason, it may be useful to focus on more proximal outcomes of working long hours that serve as risk factors for CHD. These factors would include risks for metabolic syndrome (high blood pressure, high cholesterol, high triglycerides). In sum, we make the following recommendations:

- Investigators should stop creating arbitrary categories of work hours and rely on continuous measures.
- Investigators should assess the consistency and duration of exposure to long work hours.

Confounding Variables Many of the better prospective cohort studies control for a broad array of potentially confounding variables. Age, sex, education, and SES are often among the variables that are typically included as control variables in epidemiological studies. In addition, some studies control for health-related behaviors (e.g., smoking, alcohol consumption, and lack of exercise) as well as CHD risk factors (e.g., BMI, cholesterol, and blood pressure). In terms of causal inference, important variables to control for are those that correlate not just with the outcomes of interest, but also with working hours. Partialling out the effects of variables that correlate only with the outcome can remove extraneous variance, and thus provide more precision in statistical tests by reducing the size of the error term, but it does not address the problem of endogeneity. For this reason, it is useful to know what variables tend to correlate with the number of working hours. Among the demographic predictors of working hours are sex (men higher), age (older workers work fewer, at least for men), occupation (white collar higher than clerical or blue collar), shiftwork (non-standard shifts higher), and self-employment status (self-employed higher) (Shields 1999). These are clearly important variables to control for as they tend to be significant predictors of health outcomes as well.

SES is also an interesting variable and worthy to consider in future research, because it appears to have a suppressor effect on the relationship between working hours and health (Virtanen et al. 2012b). Although high SES workers tend to work more hours, high SES is also associated with generally better health outcomes. Further, SES is related to other variables that may be independent risk factors for long working hours, poor health, or both. Thus it may be important to control for SES in order to avoid *underestimating* the impact of working hours. At least one large study (Kivimaki et al. 2015b) also found an interaction effect with SES, with long work hours posing a significant risk for stroke only for those in the lowest tertile SES group. These results suggest that SES may not be merely a confounding variable, but a moderator variable as well.

Ng and Feldman's (2008) meta-analysis highlighted other predictors of long working hours that investigators need to consider. Some of the important ones are work-related variables. Those that seem to have a protective effect, in that they are positively correlated with both work hours and many

health outcomes, are job autonomy ($r = .09$), opportunities for learning ($r = .14$), and job level ($r = .22$). Like SES, these variables might act as suppressors and should be included in estimates of work hour effects. Other work-related factors, however, are positively associated with work hours and are also predictors of poor mental and physical health. The most important of these are pressures for performance ($r = .31$), job demands ($r = .41$), and role overload ($r = .25$). Failure to control for such variables can generate spuriously significant relationships between long working hours and health outcomes ranging from depression to CHD.

Another issue worthy of consideration is the practice of starting prospective cohort studies with individuals who are free from any symptoms (CHD, depression, etc.). Whereas this is a common practice in epidemiology, and helps rule out reverse causality explanations (i.e., that unhealthy people work fewer hours as a result of their poor health), it might also cause investigators to overlook significant impacts of long work hours on the exacerbation of existing disorders. Working long hours might not have a large impact on the development of a health condition, such as CHD or depression, but it might make such conditions worse. Thus prior health status is a complex factor, and investigators need to be thoughtful in how they analyze it. Our suggestion would be to include both healthy and unhealthy individuals in the baseline cohort of a prospective design. If the design is truly longitudinal—that is, work hours were measured at multiple time points throughout the prospective period—then prior health status could be evaluated as a predictor of subsequent work hours, a control for the effects of work hours on the incidence of new cases, and as a moderator variable. In the last case, long work hours might be found to exacerbate health complaints for those who already experience them but have little effect on those not so afflicted.

Finally, investigators often control for a class of variables that are generally recognized as risk factors for certain health outcomes, especially CHD and stroke. Most frequently controlled covariates are smoking, alcohol consumption, and BMI. Some studies additionally controlled for cholesterol, blood pressure, and diabetes (e.g., Kivimaki et al. 2015a, b). With these control variables, there is a danger of *overcontrolling* if the putative confounding variable actually operates as an intervening variable. Virtually all of these variables are plausible outcomes of long working hours and also represent risk factors for disease outcomes such as CHD and stroke. Studies have reported significant associations of long work hours with weight gain (for men), smoking frequency, and drinking (for women) (Shields 1999), establishing them as plausible mediators. Diminished sleep is also a plausible mediator, as it has been linked to very long work shifts (Lockley et al. 2004), and was reported as an independent predictor of CHD along with long work hours (Cheng et al. 2014). Investigators must carefully consider the theoretical arguments regarding mediating pathways between long work hours and health outcomes as they consider adding variables to their list of controls. In summary, we make the following recommendation:

- When conducting research on the effects of work hours, researchers should also measure and control for established confounding variables, but they need to be careful to recognize potential mediating variables and not treat them as nuisance variables to be controlled.

Levels of Analysis Issues One issue that has received scant attention in this literature concerns levels of analysis. Virtually all studies investigating effects of long work hours on health outcomes do so at the individual level of analysis. This is not surprising because working hours are easily measured at the individual level and most hypothesized mediating pathways (e.g., fatigue, physiological responses to stress, poor health behaviors) are also individual-level factors. People work in a context, however, and Bliese et al. (2002) have argued that the work hours of whole groups can be expected to reflect requirements for them to do so, whereas individuals are often free to choose their hours and may do so for various reasons such as increased pay and promotion potential. The plausibility of this argument is buttressed by meta-analytic findings reported by Ng and Feldman (2008) who found that salary ($r = .24$), promotion ($r = .15$) and career satisfaction ($r = .15$) are all positively related to long work hours. One cannot rule out the explanation that such outcomes are actually the result of long working hours, but the expectation of receiving such benefits is certainly a plausible motivational antecedent. Individuals may also choose to work long hours because of their general orientation toward work and their careers, as suggested by a meta-analytic correlation ($r = .25$) between work hours and work centrality values (Ng and Feldman 2008).

A striking illustration of how level effects may influence the size of the relationship between work hours and well-being is a study by Bliese and Halverson (1996), who surveyed 7238 US Army soldiers in non-combat situations about daily working hours and psychological well-being (depression, anxiety, somatic complaints, positive well-being, and emotional control). Bliese and Halverson found that the individual-level correlation was significant but small ($r = -.16$), which is typical of the effect size in this literature. In contrast, however, the correlation between aggregated work hours and aggregated well-being for the 99 army companies was $r = -.71$. Moreover, estimates of within-group agreement (ICC(1) = .17, ICC(2) = .92) indicated that about 17% of the variance in work hours was attributed to company affiliation. Likewise, Schmitt et al. (1980) reported a similar finding for occupational groups.

The finding that group-level correlations are much higher than individual-level correlations does not necessarily imply an effect of a lack of volition on the choice of workers, for other factors are likely to vary across work groups under different commanders as well across occupational groups. Any such group or occupational differences might be confounded with working hours and produce a spurious finding. Nevertheless, this is a phenomenon that deserves further exploration. If it can be shown, by systematically controlling for group differences on other factors that affect well-being, that there is a strong mandated work hours effect, this would have clear implications for policy-making, which motivates much of the work hours and health research.

Future Directions

Our review indicates that the relationship between work hours and well-being is weaker and less consistent than expected, given the conventional wisdom that working long hours is detrimental to employee health and well-being. Possible explanations for these modest findings are that (a) the theory (i.e., that working hours represents a demand that elicits a stress response) is wrong, (b) research has not sufficiently considered boundary conditions that may mitigate the effects of working long hours on well-being, and/or (c) methodological issues have obscured the true nature of the relationship between working hours and well-being (Kluger and Tikochinsky 2001). In the section that follows, we unpack these issues and offer directions for future research.

Inadequate Theory In general, prior research has adopted a stressor-strain framework when considering how working hours might impact well-being. This framework suggests that long working hours is a job demand that has the potential to elicit a stress response. To date, however, research has not fully integrated working hours into a stressor-strain framework by explicitly considering the processes or mechanisms through which it might cause employees to experience strain and, subsequently, more negative psychological, behavioral, and physical health outcomes.

Like other job demands, working long hours is likely to impact recovery, depletion, and self-regulation, both in the short and long term. As such, it would behoove scholars to further consider working hours within the context of a dynamic theory of self-regulation, such as ego depletion theory (Baumeister et al. 1998), which explains how work demands have the potential to draw down intrapersonal resources that are necessary for optimal functioning. Being a resource-based framework, ego depletion theory suggests that people have limited resources (e.g., attention and energy) that are used to control behavior (Baumeister et al. 1998). These resources allow people to move toward goal states, while also suppressing impulses and resisting temptation. Employees expend

these resources throughout the day and there is evidence that when they are depleted, employees experience negative psychological (e.g., depressed mood, burnout, and fatigue: Goldberg and Grandey 2007; Webb and Sheeran 2003), physiological (e.g., autonomic reactivity: Gendolla and Richter 2005; Wright et al. 2007), and behavioral outcomes (deviant behaviors, job disengagement: Christian and Ellis 2011; Rosen et al. 2016; Lanaj et al. 2014). Off-the-job activities can, however, lead to recovery of depleted resources and mitigate these effects. Moreover, resource depletion processes are cumulative and dynamic, such that they may occur over days, weeks, months, or years. The key point is that insufficient recovery is associated with diminished well-being and that working long hours has the potential to draw down resources while simultaneously inhibiting sufficient recovery, making it somewhat of a double-edged sword.

Another relevant framework that expanded upon a stressor-strain model is the stressor-detachment model proposed by Sonnentag (2010). Sonnentag and Fritz (2015) reported that there is some support for psychological detachment as either a mediator or moderator of the relation between work-related stressors and worker well-being. To the extent that longer work hours results in less time away from work for recovery, workers have less time to psychologically detach from work which may be detrimental to well-being (Sanz-Vergel et al. 2011). In sum, we make the following recommendation:

- Work hours research needs to be better integrated into a stressor-strain framework that better identifies processes through which work hours cause negative health outcomes.

Boundary Conditions Characteristics of individual employees, work, and the context in which work is embedded are known to influence the extent to which job demands are depleting. In the context of working hours, however, such mitigating factors have not been systematically investigated or fully integrated into research examining the relationship between long working hours and well-being, despite evidence that factors such as work hour control and flexibility, nonwork demands, age, gender, job complexity, job demands, and norms have the potential to amplify or attenuate the effects of long working hours on health outcomes. These moderating factors are discussed below.

Ego depletion theory suggests that autonomy and control serve as individual coping resources that mitigate the extent to which demands from work are ultimately depleting (Moller et al. 2006; Muraven et al. 2008; Prem et al. 2016). Consistent with this perspective, research indicates that a variety of control and autonomy-related variables impact the relationship between working long hours and well-being. For example, owing to a number of personal and professional reasons,

individuals have different preferences for the number of hours they work, with some preferring to work more hours and others preferring to work fewer. A recent study by Sturman and Walsh (2014) indicates that work hours fit (i.e., the difference between the actual number of hours an employee works and the desired number of hours worked by that employee) has an impact on outcomes of hours worked, such that employees experience more positive outcomes (e.g., reduced medical absences, greater life satisfaction) when there is congruence between preferred and actual hours worked. There is also evidence that control over working hours and schedule autonomy have the potential to weaken the linkage between hours worked and negative health outcomes (Ala-Mursala et al. 2006; Tucker and Rutherford 2005). In addition, flextime schedules, which give employees a degree of control over their work hours and facilitate management of competing work and non-work demands, also appear to have benefits for employees, including increased job satisfaction and reduced absenteeism (Baltes et al. 1999). These findings around work hour control, autonomy, and flextime are consistent with Sturman and Walsh (2014), given that employees who have greater control over their work hours are more likely to work hours that are congruent with their preferences. Thus, providing employees with some control over the number of hours they work, and when they work those hours, has the potential to play an important role in mitigating the effects of long working hours on well-being. However, it still needs to be empirically verified that having control over one's work hours can significantly moderate their effects on well-being.

Another topic that has received attention alongside time spent doing paid work (i.e., working hours) is time spent performing domestic tasks (i.e., "household work"). In general, research indicates that domestic work is detrimental to the well-being of individuals who also perform paid work, having a positive relationship with insomnia, exhaustion, and depression (e.g., Tierney et al. 1990). This phenomenon is known as the "second shift" (Hochschild and Machung 2012). These findings make sense, given that there is likely little time left for recovery after work for individuals who devote a significant amount of time to completing domestic tasks (e.g., taking care of children, preparing meals, doing laundry) during the time that they spend away from work. Interestingly, perceived inequity in the household distribution of labor amplifies these effects for women, given that women have traditionally devoted more time to domestic duties than men (Glass and Fujimoto 1994; Cleveland et al. 2015). There is also evidence that the combination of time spent on domestic work and paid work has an "enhanced effect over paid work alone particularly for women" (Hunt and Annandale 1993, p. 659). Previously discussed factors such as autonomy and control over work hours are likely to have an impact on the ability of both men and women to balance work and non-work demands, which may allow individuals to include more time for

recovery in their schedules. Therefore, we suggest that future studies further consider how work and non-work demands combine with other factors (e.g., work hour autonomy and control, flextime) to mitigate the effects of long working hours on well-being. A deeper exploration of gender effects (discussed below) is also warranted, given that gender roles at home and work have continued to evolve over the past quarter century.

With regard to demographics, Ng and Feldman (2008) included a number of exploratory moderator analyses in their meta-analysis that considered the relationship between work hours and well-being. Their results indicated that age and gender have the potential to impact the relationship between work hours and well-being. In particular, work hours demonstrate stronger positive relationships with mental strain and physical health symptoms for men relative to women. This finding was consistent with Ng and Feldman's (2008) expectations that men might be more vulnerable to the effects of workplace demands and stressors because they are more concerned about their careers. Results further indicated that the positive relationship between work hours and physical health symptoms was weaker for older workers than for younger ones, which also was consistent with the authors' expectation that salient organizational identity has a stronger relationship with work hours for younger workers. Given the exploratory nature of these analyses, however, further research is needed to understand when and why men and younger workers are may be more vulnerable to the effects of long work hours. In addition, these findings should be considered in tandem with the research which has demonstrated that hours spent on non-paid work hours (e.g., domestic or household activities) exert a negative influence on well-being, given the changing nature of gender roles over the past quarter century.

In addition to demographics, characteristics of work itself (e.g., job complexity and job demands) are also likely to determine the extent to which long work hours are depleting and, hence, detrimental to well-being. Jobs that are more complex require greater skill variety, more skills and abilities, and intensive cognitive effort. At the same time, more complex jobs often allow employees to have greater autonomy and discretion with regard to their work. These desirable work characteristics "promote the development of vital resources, including skills and psychological resources, for managing work and family demands" (Valcour 2007, p. 1515). For these reasons, it has been argued that job complexity ameliorates the negative effects of work and working long hours on outcomes such as work/family conflict and job dissatisfaction (Valcour 2007). Research has not, however, provided consistent evidence for this effect (Ng and Feldman 2008; Valcour 2007). Moreover, an alternative hypothesis is that employees who work longer hours on more complex jobs will be more depleted because they must expend more cognitive resources doing their work compared with those who work long hours on less complex

work. These competing hypotheses should be investigated in future empirical studies. Moreover, in order to fully consider the effects of job complexity on the relationship between work hours and well-being, scholars should more directly assess how job complexity affects the relationship between long work hours and the mediating mechanism of cognitive resource depletion.

Similar to job complexity and consistent with the stressor-detachment model (Sonnentag 2010), long work hours may be more depleting to employees who have excessive job demands (e.g., pressure to work fast, limited time to complete work), given that work hours increase the amount of exposure that employees have to those demands. As such, employees in more demanding jobs will have resources drawn down at a greater rate and will, therefore, have a greater need for recovery off the job. Unfortunately, this might create a dangerous cycle because employees who work more hours will have less time available to engage in off-the-job recovery activities. As such, future research should consider how such cycles of depletion may develop as a result of working more hours in highly demanding jobs and researchers should investigate the efficacy of interventions aimed at curbing these cycles.

Finally, our discussion of levels of analysis issues suggests that the context in which work is embedded may have an impact on how employees experience long working hours. Taking a systems perspective, context can be viewed as local (e.g., work group or organizational), familial, or more broadly (e.g., regional or national culture). For local context, it is important to consider work-related norms around hours worked and how individuals who meet, exceed, or fall below those norms are affected by their work hours. Employees who are unable to meet organization or work group norms may experience anxiety around not doing what is expected of them. In addition, these individuals may be ostracized or alienated by group members for not contributing in a way that is expected. Likewise, individuals who exceed work group norms may also experience negative treatment from the workgroup, as they might be perceived as rate busters. With regard to the family context, advances in work/family research have examined dyadic relationships among spouses/partners such that the effect of work hours on well-being may also be related to employees' family expectations and experiences. For example, Wayne et al. (2013) theorized and found empirical support for the notion that a spouse/partner's attitude toward the employee's job or employing organization may be affected by the spouse/partner's perceptions of the employee's work hours.

In terms of regional and national context, these norms are more cultural and the implications are that those whose work hours fall below or above what is expected in a given society may be viewed negatively by those around them. Research also suggests that social context surrounding work serves to explain the broader meaning of time spent working (Kleiner et al. 2015). For example, in countries where the welfare state is stronger, there is less dependence on work for outcomes such as healthcare and retirement benefits (Bambra and Eikemo 2009; Bell and Freeman 2001). As such, work is viewed as less central to life in these contexts and time spent working undesirable hours has a weaker relationship with higher order, long-term goals, suggesting that workers in some cultures will prefer to work fewer hours than workers in other cultures. Recent studies (e.g., Kleiner et al. 2015) have, in fact, supported this perspective, indicating that in European countries such as Germany, which traditionally has had low income inequality and a strong welfare state, long working hours have unequivocal negative effects on mental health. In contrast, research has consistently shown that the same relationships are more inconsistent and complex in countries such as the USA, where time spent working is less regulated and the government has traditionally been less supportive of benefits around health care and retirement, to name a few. Interestingly, Kleiner et al.'s (2015) study also demonstrated that in the USA, employees who work fewer hours (i.e., part-time) report worse mental health, a finding that was not replicated in the authors' German sample. There are likely a number of factors that explain this relationship. In particular, in the USA, part-time employees tend to receive lower wages and fewer benefits, while employees working fewer hours in Germany still receive a full complement of what are viewed as work benefits in the USA (e.g., health insurance, job protection, vacation time), which may serve to mitigate the effects of working fewer hours (Kleiner et al. 2015).

Further supporting the notion that cultural differences and norms might impact how long work hours impact well-being, a study by Spector et al. (2004) indicated that there are differences in how people from three culturally distinct regions—Anglo (Australia, Canada, England, New Zealand, and USA), China (Hong Kong, People's Republic of China, and Taiwan), and Latin America (Argentina, Brazil, Colombia, Ecuador, Mexico, Peru, and Uruguay) experience long work hours. Their results indicated that the positive relation between work hours and work/family stressors (e.g., work/family pressure) was strongest for Anglos. Spector et al. (2004) interpreted these results as supporting their theory, which suggests that for Anglos, working extra hours is viewed as taking time away from family activities. In China and Latin America, however, this is not the case because in those regions working long hours is viewed differently by employees and their families.

The study of work and cultural norms is, obviously, a complex area of research, and more studies are necessary to determine how, when, and why norms at different levels have an impact on the relationship between time spent working and various indicators of well-being. Further complicating the issue, norms at different levels (e.g., organizational and cultural) can come into conflict. For example, in Japan, it is acceptable to work very long hours each week. This is in stark contrast to most European countries, where the work week is typically

shorter. A Japanese firm opening an office in Europe would need to anticipate that its employees would not respond favorably to working the same hours as Japanese workers. Likewise, an executive from a Swedish firm should anticipate working longer hours if she were to take a similar position in the USA, given the cultural differences that exist around long working hours in both countries. We are optimistic that future empirical study will take advantage of these research opportunities and consider how group, organizational, and cultural norms combine to determine how work hours affect well-being across cultures. Our summary recommendation is thus:

- Future studies should systematically examine characteristics of individual employees and the work context that mitigate or exacerbate the effects of work hours on health outcomes.

Conclusion

Conventional wisdom suggests that long working hours have a negative effect on well-being. Research conducted over the past few decades, however, has failed to provide clear and consistent evidence for such an effect. Drawing from our review of the literature, we identified a set of research opportunities for future scholars interested in shedding additional light on this topic. Our recommendations include applying more nuanced theoretical frameworks to explain relations between long work hours and well-being, focusing on boundary conditions of the effects of long work hours, and addressing methodological issue that have emerged in this literature. Studies that tackle these issues will provide an avenue for research in the organizational sciences to better inform public policy and professional practice going forward.

References

Ala-Mursala, L., Vahtera, J., Kouvonen, A., Vaananen, A., Linna, A., Pentti, J., & Kivimaki, M. (2006). Long hours in paid and domestic work and subsequent sickness absence: does control over daily working hours matter? *Occupational and Environmental Medicine, 63*, 608–616.

Amagasa, T., & Nakayama, T. (2012). Relationship between long working hours and depression in two working populations: a structural equation approach. *Journal of Occupational and Environmental Medicine, 54*, 868–874.

Baltes, B. B., Briggs, T. E., Huff, J. W., Wright, J. A., & Neuman, G. A. (1999). Felxible and compressed workweek schedules: a meta-analysis of their effects on work-related criteria. *Journal of Applied Psychology, 84*, 496–513.

Bambra, C., & Eikemo, T. A. (2009). Welfare state regimes, unemployment, and health: a comparative study of the relationship between unemployment and self-reported health in 23 European countries. *Journal of Epidemiology and Community Health, 63*, 92–98.

Bannai, A., & Tamakoshi, A. (2014). RThe association between long working hours and health: a systematic review of epidemiological evidence. *Scandinavian Journal of Work, Environment, and Health, 40*, 5–18.

Barnett, R. C. (2006). Relationship of the number and distribution of work hours to health and quality-of-life (QOL) outcomes. In P. L. Perrewe' & D. C. Ganster (Eds.), *Research in occupational stress and well-being, volume 5: employee health, coping, and methodologies* (pp. 99–138). Amsterdam: Elsevier.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *Journal of Personality and Social Psychology, 74*(5), 1252–1265.

Bell, L. A., & Freeman, R. B. (2001). The incentive for working hard: explaining hours worked differences in the US and Germany. *Labour Economics, 8*(2), 181–202.

BBC. http://www.bbc.com/news/business-34677949. Accessed 26 August 2016.

Bliese, P. D., & Halverson, R. R. (1996). Individual and nomothetic models of job stress: an examination of work hours, cohesion, and well-being. *Journal of Applied Social Psychology, 26*, 1171–1189.

Bliese, P. M., Jex, S. M., & Halverson, R. R. (2002). Integrating multi-level analyses and occupational stress theory. In P. Perrewe & D. Ganster (Eds.), *Research in occupational stress and well-being, Vol. 2: historical and current perspectives on stress and health* (pp. 217–259). Amsterdam: Elsevier.

Bonett, D. G. (2007). Transforming odds ratios into correlations for meta-analytic research. *American Psychologist, 62*, 254–255.

Bosma, H., Marmot, M., Hemingway, H., Nicholson, A. C., Brunner, E., & Stansfeld, S. A. 1997. Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study. British Medical Journal, 314, 558–565. https://www.cdc.gov/chronicdisease/resources/publications/aag/diabetes.htm. Accessed 25 August 2016.

Caruso, C. C., Bushnell, T., Eggerth, D., Heitmann, A., Kojola, B., Newman, K., Rosa, R., Sauter, S., & Vila, B. (2006). Long working hours, safety, and health: toward a national research agenda. *American Journal of Industrial Medicine, 49*, 930–942.

Cheng, Y., Du, C.-L., Hwang, J.-J., Chen, I.-S., Chen, M.-F., & Su, T.-C. (2014). Working hours, sleep duration and the risk of acute coronary heart disease: a case-control study of middle-aged men in Taiwan. *International Journal of Cardiology, 171*, 419–422.

Christian, M. S., & Ellis, A. P. (2011). Examining the effects of sleep deprivation on workplace deviance: a self-regulatory perspective. *Academy of Management Journal, 54*(5), 913–934.

Cleveland, J. N., Fisher, G. G., & Sawyer, K. (2015). Work-family equality: importance of a level playing field at home. In M. J. Mills (Ed.), *Gender and the work/family experience: an intersection of two domains* (pp. 179–201). New York, NY: Springer.

Digby, P. G. N. (1983). Approximating the tetrachoric correlation coefficient. *Biometrics, 39*, 753–757.

Driesen, K., Jansen, N. W. H., Kant, I., Mohren, D. C. L., & van Amelsvoort, L. G. P. M. (2010). Depressed mood in the working population: associations with work schedules and working hours. *Chronobiology International, 27*, 1062–1079.

Eaton, W. W., Martins, S. S., Nestadt, G., Bienvenu, O. J., Clarke, D., et al. (2008). The burden of mental disorders. *Epidemiologic Reviews, 30*, 1–14.

Fritz, C., & Sonnentag, S. (2006). Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences. *Journal of Applied Psychology, 91*(4), 936–945.

Ganster, D. C. (2005). Executive job demands: suggestions from a stress and decision making perspective. *Academy of Management Review, 30*, 492–502.

Ganster, D. C., & Rosen, C. R. (2013). Work stress and employee health: a multidisciplinary review. *Journal of Management, 39*, 1085–1122.

Gendolla, G. H., & Richter, M. (2005). Ego involvement and effort: cardiovascular, electrodermal, and performance effects. *Psychophysiology, 42*(5), 595–603.

Glass, J., & Fujimoto, T. (1994). Housework, paid work, and depression among husbands and wives. *Journal of Health and Social Behavior, 35*, 179–191.

Goh, J., Pfeffer, J., & Zenios, S. A. (2016). The relationship between workplace stressors and mortality and health costs in the United States. *Management Science, 62*, 608–628.

Goldberg, L. S., & Grandey, A. A. (2007). Display rules versus display autonomy: emotion regulation, emotional exhaustion, and task performance in a call center simulation. *Journal of Occupational Health Psychology, 12*(3), 301–318.

Greenhaus, J. H., & Beutell, N. J. (1985). Sources of conflict between work and family roles. *Academy of Management Review, 10*(1), 76–88.

Hochschild, A., & Machung, A. (2012). *The second shift: working families and the revolution at home.* New York, NY: Penguin Books.

Hunt, K., & Annandale, E. (1993). Just the job? Is the relationship between health and domestic and paid work gender-specific? *Sociology of Health and Illness, 15*, 632–664.

Johnson, R. E., Rosen, C. C., & Djurdjevic, E. (2011). Assessing the impact of common method variance on higher-order multidimensional constructs. *Journal of Applied Psychology, 96*, 744–761.

Kang, M.-Y., Park, H., Seo, J.-C., Lim, D., Lim, Y.-H., Lim, S., Cho, S.-H., & Hong, Y.-C. (2012). Long working hours and cardiovascular disease. *Journal of Occupational and Environmental Medicine, 54*, 532–537.

Kivimaki, M., Jokela, M., Nyberg, S. T., Singh-Manoux, A., Fransson, E., Alfredsson, L., Bjomer, J., Borritz, M., Burr, H., Casini, A., Clays, E., De Bacquer, D., Dragano, N., Erbel, R., Geuskens, G. A., Hamer, M., Hooftman, W. E., Houtman, I. L., Jockel, K.-H., Kittel, F., Knutsson, A., Koskenvuo, M., Lunau, T., Madsen, I. E. H., Nielsen, M. L., Nordin, M., Oksanen, T., Pejtersen, J. H., Pentti, J., Rugulies, R., Salo, P., Shipley, M. J., Siegrist, J., Steptoe, A., Suominen, S. B., Theorell, T., Vahtera, J., Westerholm, P. J. M., Westerlund, H., O'Reilly, D., Kumari, M., Batty, G. D., Ferrie, J. E., & Virtanen, M. (2015a). Long working hours and risk of coronary heart disease and stroke: a systematic review and meta-analysis of published and unpublished data for 603 838 individuals. *The Lancet, 140-6736*(15), 60295–60291 http://dx.doi.org/10.1016/SO.

Kivimaki, M., Virtanen, M., Kawachi, I., Nyberg, S. T., Alfredsson, L., Batty, G. D., Bjorner, J. B., Borritz, M., Brunner, E. J., Burr, H., Dragano, N., Ferrie, J. E., Fransson, E. I., Hamer, M., Heikkila, K., Knutsson, A., Koskenvuo, M., Madsen, I. E., Nielsen, M. L., Nordin, M., Oksanen, T., Pejtersen, J. H., Pentti, J., Rugulies, R., Salo, P., Siegrist, J., Steptoe, A., Suominen, S., Theorell, T., Vahtera, J., Westerholm, P. J. M., Westerlund, H., Singh-Manoux, A., & Jokela, M. (2015b). Long working hours, socioeconomic status, and the risk of incident type 2 diabetes: a meta-analysis of published and unpublished data from 222 120 individuals. *Lancet Diabetes Endocrinology, 3*, 27–34.

Kleiner, S., Schunck, R., & Schomann, K. (2015). Different contexts, different effects? Work time and mental health in the United States and Germany. *Journal of Health and Social Behavior, 56*, 98–113.

Kossek, E. E., Ruderman, M. N., Braddy, P. W., & Hannum, K. M. (2012). Work–nonwork boundary management profiles: a person-centered approach. *Journal of Vocational Behavior, 81*(1), 112–128.

Kluger, A. N., & Tikochinsky, J. (2001). The error of accepting the "theoretical" null hypothesis: the rise, fall, and resurrection of commonsense hypotheses in psychology. *Psychological Bulletin, 127*(3), 408–423.

Lanaj, K., Johnson, R. E., & Barnes, C. M. (2014). Beginning the work-day yet already depleted? Consequences of late-night smartphone use and sleep. *Organizational Behavior and Human Decision Processes, 124*(1), 11–23.

Landrigan, C. P., Fahrenkopf, A. M., Lewin, D., Sharek, P. J., Barger, L. K., Eisner, M., & Sectish, T. C. (2008). Effects of the accreditation council for graduate medical education duty hour limits on sleep, work hours, and safety. *Pediatrics, 122*(2), 250–258.

Lockley, S. W., Cronin, J. W., Evans, E. E., Cade, B. E., Lee, C. J., Landrigan, C. P., Rothschild, J. M., Katz, J. T., Lilly, C. M., Stone, P. H., Aeschbach, D., & Czeisler, C. A. (2004). Effect of reducing interns' weekly work hours on sleep and attentional failures. *The New England Journal of Medicine, 351*, 1829–1837.

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: the moderating role of autonomy. *Personality and Social Psychology Bulletin, 32*(8), 1024–1036.

Muraven, M., Gagné, M., & Rosman, H. (2008). Helpful self-control: autonomy support, vitality, and depletion. *Journal of Experimental Social Psychology, 44*(3), 573–585.

Nagashima, S., Suwazono, Y., Okubo, Y., Uetani, M., Kobayashi, E., Kido, T., & Nogawa, K. (2007). Working hours and mental and physical fatigue in Japanese workers. *Occupational Medicine, 57*, 449–452.

Ng, T. W. H., & Feldman, D. C. (2008). Long work hours: a social identity perspective on meta-analysis data. *Journal of Organizational Behavior, 29*, 853–880.

Prem, R., Kubicek, B., Diestel, S., & Korunka, C. (2016). Regulatory job stressors and their within-person relationships with ego depletion: the roles of state anxiety, self-control effort, and job autonomy. *Journal of Vocational Behavior, 92*, 22–32.

Robinson, J. P., Chenu, A., & Alvarez, A. S. (2002). Measuring the complexity of hours at work: the weekly work grid. *Monthly Labor Review, 125*, 44–54.

Rosen, C. C., Koopman, J., Gabriel, A., & Johnson, R. E. (2016). Who strikes back? A daily investigation of when and why incivility begets incivility. *Journal of Applied Psychology.*

Sanz-Vergel, A. I., Demerouti, E., Bakker, A. B., & Moreno-Jiménez, B. (2011). Daily detachment from work and home: the moderating effect of role salience. *Human Relations, 64*, 775–799.

Schmitt, N., Colligan, M. J., & Fitzgerald, M. (1980). Unexplained physical symptoms in eight organizations: individual and organizational analyses. *Journal of Occupational Psychology, 53*, 305–317.

Shields, M. (1999). Long working hours and health. *Health Reports, 11*, 33–48.

Solovieva, S., Lallluka, T., Virtanen, M., & Viikar-Juntura, E. (2013). Psychosocial factors at work, long work hours, and obesity: a systematic review. *Scandinavian Journal of Work, Environment, and Health, 39*, 241–258.

Sonnentag, S. (2010). Recovery from fatigue: the role of psychological detachment. In P. L. Ackerman (Ed.), *Cognitive fatigue: the current status and future for research and application* (pp. 253–272). Washington, DC: American Psychological Association. doi:10.1037/12343-012.

Sonnentag, S., & Fritz, C. (2015). Recovery from job stress: the stressor-detachment model as an integrative framework. *Journal of Organizational Behavior, 36*(S1), S72–S103.

Sparks, K., Cooper, C., Fried, Y., & Shirom, A. (1997). The effects of work hours on health: a meta-analytic review. *Journal of Occupational and Organizational Psychology, 70*, 391–408.

Spector, P. E., Cooper, C. L., Poelmans, S., Allen, T. D., O'Driscoll, M., Sanchez, J. I., Siu, O. L., Dewew, P., Hart, P., Lu, L., De Moraes, L. F. R., Ostrognay, G. M., Sparks, K., Wog, P., & Yu, S. (2004). A cross-national comparative study of work-family stressors, working

hours, and well-being: China and Latin America versus the Anglo world. *Personnel Psychology, 57,* 119–142.

Spector, P. E., Dwyer, D. J., & Jex, S. M. (1988). Relation of job stressors to affective, health, and performance outcomes: a comparison of multiple data sources. *Journal of Applied Psychology, 73,* 11–19.

Steptoe, A., Wardle, J., Lipsey, Z., Mills, R., Oliver, G., Jarvis, M., & Kirschbaum, C. (1998). A longitudinal study of work load and variations in psychological well-being, cortisol, smoking, and alcohol consumption. *Annals of Behavioral Medicine, 20*(2), 84–91.

Sturman, M. C., & Walsh, K. (2014). Strengthening the employment relationship: the effects of work-hours fit on key employee attitudes. *Journal of Organizational Behavior, 35,* 762–784.

Tarumi, K., Hagihara, A., & Morimoto, K. (2003). A prospective observation of health defects associated with working hours. *Industrial Health, 41,* 101–108.

Tarumi, K., Hagihara, A., & Morimoto, K. (2004). Moderating effects of psychological job strain on the relationship between working hours and health: an examination of white-collar workers employed by a Japanese manufacturing company. *Journal of Occupational Health, 46,* 345–351.

ten Brummelhuis, L.L., Rothbard, N.P., & Uhrich, B. (2016) Beyond nine to five: Is working to excess bad for health? Academy of Management Discoveries.

Tierney, D., Romito, P., & Messing, K. (1990). She ate not the bread of idleness: exhaustion is related to domestic and salaried working conditions among 539 Québec hospital workers. *Women & Health, 16*(1), 21–42.

Tucker, P., & Rutherford, C. (2005). Moderators of the relationship between long work hours and health. *Journal of Occupational Health Psychology, 10,* 465–476.

Valcour, M. (2007). Work-based resources as moderators of the relationship between work hours and satisfaction with work-family balance. *Journal of Applied Psychology, 92,* 1512–1523.

Virtanen, M., Ferrie, J. E., Singh-Manoux, A., Shipley, M. J., Stansfeld, S. S., Marmot, M. G., Ahola, K., Vahtera, J., & Kivimaki, M. (2011). *Psychological Medicine, 41,* 2485–2494.

Virtanen, M., Heikkela, K., Jokela, M., Ferrie, J. E., Batty, G. D., Vahtera, J., & Kivimaki, M. (2012a). Long working hours and coronary heart disease: a systematic review and meta-analysis. *American Journal of Epidemiology, 176,* 586–596.

Virtanen, M., Stansfeld, S. A., Fuhrer, R., Ferrie, J. E., & Kivimaki, M. (2012b). Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. *PloS One, 7*(1), e30719.

Wayne, J. H., Casper, W. J., Matthews, R. A., & Allen, T. D. (2013). Family-supportive organization perceptions and organizational commitment: the mediating role of work–family conflict and enrichment and partner attitudes. *Journal of Applied Psychology, 98*(4), 606–622.

Webb, T. L., & Sheeran, P. (2003). Can implementation intentions help to overcome ego-depletion? *Journal of Experimental Social Psychology, 39*(3), 279–286.

West, C. P., Tan, A. D., Habermann, T. M., Sloan, J. A., & Shanafelt, T. D. (2009). Association of resident fatigue and distress with perceived medical errors. *JAMA, 302*(12), 1294–1300.

World Health Organization. (1994). *Composite International Diagnostic Interview (CIDI): Researcher's Manual.* American Psychiatric Publications Inc.

Wright, R. A., Junious, T. R., Neal, C., Avello, A., Graham, C., Herrmann, L., & Walton, N. (2007). Mental fatigue influence on effort-related cardiovascular response: difficulty effects and extension across cognitive performance domains. *Motivation and Emotion, 31*(3), 219–231.