Systematic review on the association between employee worktime control and work–non-work balance, health and well-being, and job-related outcomes

by Hylco H Nijp, MSc,1 Debby GJ Beckers, PhD,1 Sabine AE Geurts, PhD,1 Philip Tucker, PhD,2,3 Michiel AJ Kompier, PhD1


Objectives The aim of this review was to assess systematically the empirical evidence for associations between employee worktime control (WTC) and work–non-work balance, health/well-being, and job-related outcomes (eg, job satisfaction, job performance).

Method A systematic search of empirical studies published between 1995–2011 resulted in 63 relevant papers from 53 studies. Five different categories of WTC measurements were distinguished (global WTC, multidimensional WTC, flextime, leave control, and “other subdimensions of WTC”). For each WTC category, we examined the strength of evidence for an association with (i) work–non-work balance, (ii) health/well-being, and (iii) job-related outcomes. We distinguished between cross-sectional, longitudinal, and intervention studies. Evidence strength was assessed based on the number of studies and their convergence in terms of study findings.

Results (Moderately) strong cross-sectional evidence was found for positive associations between global WTC and both work–non-work balance and job-related outcomes, whereas no consistent evidence was found regarding health/well-being. Intervention studies on global WTC found moderately strong evidence for a positive causal association with work–non-work balance and no or insufficient evidence for health/well-being and job-related outcomes. Limited to moderately strong cross-sectional evidence was found for positive associations between multidimensional WTC and our outcome categories. Moderately strong cross-sectional evidence was found for positive associations between flextime and all outcome categories. The lack of intervention or longitudinal studies restricts clear causal inferences.

Conclusions This review has shown that there are theoretical and empirical reasons to view WTC as a promising tool for the maintenance of employees’ work–non-work balance, health and well-being, and job-related outcomes. At the same time, however, the current state of evidence allows only very limited causal inferences to be made regarding the impact of enhanced WTC.

Key terms job motivation; job performance; stress; work–family conflict; work scheduling; worktime flexibility.

During the past decades, organizations increasingly emphasized work-related flexibility in their organizational practices (1). One type of flexibility that has become more common is “temporal flexibility”, ie, flexibility regarding working times. Initially, flexible worktime arrangements were mainly implemented for the benefit of the organization (eg, mandatory overtime work and shift-work), but over the years, attention has shifted towards flexible worktime arrangements, such as worktime control (WTC) (2), that may benefit both the organization and its employees. WTC can be defined as “an employee’s possibilities of control over the duration, position, and distribution of worktime” (3). WTC comes in many forms. Well-known subdimensions include control over (i) starting and ending times of the workday (ie, flextime), (ii) when to take a break, (iii) when to take vacation or a day off, (iv) the distribution of workdays over the work week, and (v) whether and when to work overtime.

The increasing popularity of WTC can be explained
by its assumed positive effects on employee work–non-work balance, health and well-being, and performance. For instance, self-determination theory (4) and several influential occupational health theories [eg, demand–control model (5), job characteristics model (6)] state that job autonomy – of which WTC is a specific sub-dimension – is a key factor for employee motivation, health, and performance.

At a more fundamental level, we propose two regulatory mechanisms that can explain the hypothesized favorable association between WTC and indicators of health/well-being and performance: a time-regulation mechanism, and a recovery-regulation mechanism. The first mechanism implies that WTC enables workers to align their working times with their responsibilities in private life. Due to this time-regulating quality, WTC may be an excellent buffer against (time-based) work–home interference. Research has shown that a good balance between work and home results in higher worker energy, motivation, and satisfaction (7).

WTC can also be identified as a recovery-regulation mechanism. Occupational health research has highlighted the relevance of sufficient recovery, showing that insufficient recovery is a main mechanism underlying the association between stressful work and adverse health (8). According to effort–recovery theory (9), the key determinants of the balance between effort and recovery are workload and work control. From a health-perspective, high workload may adversely influence the effort–recovery balance. Workload is to a large extent determined by the amount and complexity of work, but also by temporal aspects of work (working time arrangements), since the number and distribution of work hours determine the duration and intensity of the exposure to workload, as well as the opportunities for recovery (9). Work control can be seen as a key factor in preventing worker overload and preserving a favorable effort–recovery balance. Concerning recovery, control of the temporal aspects of work (ie, WTC) may have an especially important recuperative value (10, 11) as high individual WTC allows workers to stop working before becoming too fatigued (12). In this respect, WTC can be a means for internal recovery (ie, recovery on the job), as it allows employees to take a break when they feel the need to recover. It may also enhance external recovery (in-between working periods) as it allows workers to have control over leave days, overtime work, or starting and ending times of the workday. So in brief, it can be theorized that WTC facilitates recovery opportunities and consequently can be a buffer against high fatigue and stress, and it may also stimulate vitality, work motivation, and performance.

Empirical research is necessary to find out whether such assumptions about the favorable impact of WTC on employee work–non-work balance, health/well-being, and performance are valid. During the past decades, much scientific research has addressed the trend towards increasingly diverse and flexible working times. This has resulted in a considerable number of empirical studies exploring the effects of WTC on employee and organizational outcomes. Baltes and colleagues (13) conducted a meta-analysis of 27 intervention studies that examined the association between flextime (ie, control over starting and ending times of the workday) and various organization relevant outcomes. In line with our theoretical assumptions, their meta-analysis showed that an increase in flextime was indeed associated with positive effects on productivity, job satisfaction, satisfaction with work schedule, and absenteeism. More recently, a (partly overlapping) selection of ten intervention studies was included in a high-quality systematic review about the health effects of flexible work arrangements by Joyce and coworkers (14). As regards WTC, this review focused on health effects and included both studies on flextime and self-scheduling (freedom to schedule one’s own work hours). Five intervention studies were included and the authors tentatively concluded that employee WTC has the potential to favorably influence employee health.

Despite these promising findings, broader conclusions about the effects of WTC cannot be drawn from these reviews as both had a rather specific focus. Baltes et al (13) merely focused on the effects of flextime, and the review by Joyce and colleagues (14) included only findings on two specific dimensions of WTC. Also, both reviews focused on intervention studies and excluded studies with other study designs. Furthermore, neither review provided a comprehensive overview of the “outcomes” of WTC, with Joyce et al (14) focusing on health effects and Baltes et al (13) on a particular set of job-related outcomes. Finally, with the latter meta-analysis having been published more than a decade ago, an updated overview of studies on the association between WTC and job-related outcomes is needed.

So, whilst recognizing the value of these earlier reviews, we decided to conduct a new and broader review with: (i) a more comprehensive operationalization of WTC; (ii) studies with cross-sectional, longitudinal, and intervention designs; and (iii) a broader range of relevant outcome categories, thus extending the research focus to indicators of work–non-work balance, health/well-being, as well as job-related outcomes such as job satisfaction and job performance.

The aim of the current study was therefore to review systematically the recent empirical literature on WTC and provide a complete overview of the recent empirical evidence of the associations between WTC, on the one hand, and indicators of work–non-work balance, health/well-being, and job-related outcomes on the other. We focused on the impact of “global WTC” and specific WTC categories (eg, flextime, control over leave time
and vacations, and other subdimensions, such as control over overtime). Regarding outcome categories, we focused on outcomes that are related to the time- and recovery-regulation mechanisms mentioned above: work–non-work balance, health/well-being, and job-related outcomes.

Specifically, our research questions were: (i) How strong is the empirical evidence regarding the association between (categories of) WTC and indicators of work–non-work balance, health/well-being, and job-related outcomes? (ii) In case of significant associations between (categories of) WTC and these indicators, how strong is the empirical evidence that these associations are causal in nature?

Method

Study selection

A systematic literature search was conducted within the PsycINFO and PubMed databases, the latter also including the Medline database. We confined our search to relevant empirical English language papers published between 1995 and September 2011. We used a set of keywords related to “WTC”, for example “worktime AND control”, “working schedule AND flexibility”, “self-scheduling”, and “flextime”. A complete overview of the search terms can be found in the Appendix (available online: http://www.sjweh.fi/data_repository.php). These search terms resulted in 2000 hits. In addition, 67 references within an earlier review (14) and three references from the authors’ WTC files were checked for inclusion.

The first author scanned the abstracts of these 2070 papers. Papers that did not focus on the association between WTC (and subdimensions of WTC) and indicators of work–non-work balance, health/well-being, or job-related outcomes were excluded. This first selection round resulted in exclusion of 1829 papers, with 241 papers remaining. Three papers could not be retrieved online, and their authors did not respond to requests for fulltext papers. Next, the first three authors assessed the relevance of the remaining 238 papers, using the following inclusion criteria: (i) publication type and research purpose: empirical quantitative studies (ie, assessing statistical associations among WTC and relevant outcome variables); (ii) study design: cross-sectional studies including at least 100 respondents, intervention studies with control group and pre- and post intervention measurement, and longitudinal studies. For the two latter categories, there was no restriction on the number of participants; (iii) sample: samples consisting of healthy and working individuals. Atypical samples (eg, employees with cardiovascular disease) were excluded; (iv) relevant measure(s) of WTC: only studies with measurements of WTC that fitted our definition of WTC (“control over the duration, position, and distribution of worktime” (3, p503)) were included. If only some example items of WTC were reported, the authors of the specific paper were contacted and asked for detailed information on the complete and exact measurement of WTC (papers 15 and 16). Studies were excluded if the measure of WTC also included other elements [eg, control over work location, extended workdays (17)] or was confounded with relevant outcome measures [eg, “my schedule allows me the flexibility I need to lead a balanced lifestyle” (18)].

Based on these inclusion criteria, the three assessors individually rated each study as either relevant, irrelevant, or undecided. Their ratings converged on 98% of all papers. The 21 studies that were considered “undecided” by at least one of the assessors were discussed in depth until raters agreed upon inclusion or exclusion.

Where multiple papers report on an overlapping dataset, these papers were grouped together and treated as single studies (ie, the 10-Town Study: papers 19–24); the National Study of the Changing Workforce 1992: 25, 26; the National Study of the Changing Workforce 2002: 27, 28, 29; the European Survey on Working Conditions: 30, 31; Results Only Work Environment Study Intervention at Best Buy: 32, 33). In case multiple papers addressed the same study, the paper numbers are mentioned together, divided by a slash (eg, “19/20”).

Based on this procedure, 63 papers were included in the review, representing a total of 53 studies [see figure 1 (based on 34) for an overview of the number of included papers after each inclusion-criterion step]. A complete list of all 238 papers with information on arguments for exclusion or inclusion can be obtained from the first author.

Synthesis of evidence

As the 53 selected studies showed considerable heterogeneity in terms of measurement of WTC, outcome variables, and analyses, it was not feasible to conduct a meta-analysis (14, 35). To avoid mere “vote-counting” (36), we applied a standardized index of convergence (SIC) (37) to quantify the evidence for the assumed association between WTC, on the one hand, and work–non-work balance, health/well-being, and job-related outcomes on the other. Wielenga-Meijer et al (37) define SIC as:

\[
\frac{n[\text{positive}] - n[\text{negative}]}{n[\text{total}]}
\]

In this formula, \(n[\text{positive}]\) represents “the number of studies (examining the defined relationship) that
reported a significant positive relationship, $n_{negative}$ represents the number of studies that found a significant negative relationship, and $n_{total}$ represents the total number of studies (including studies that did not find a significant association) for the defined relationship (37, p365). SIC values thus range from -1 (all included studies show negative associations) to +1 (all included studies show positive associations between WTC and the specific outcome category). A SIC close to 0 implies that the studies examining this association reported inconsistent findings or failed to find a positive or a negative association between WTC and the outcome variable of interest (37). In short, SIC represents the degree to which findings regarding the association of WTC and a specific outcome category are consistent (degree of consistency).

The combination of these SIC values and the number of studies among which this convergence was calculated serves as a measure of strength of evidence (37, see table 1). As table 1 shows, the strength of evidence for each examined association can be either “strong” (+++ or ---), “moderately strong” (++ or --), “limited” (+ or -), “inconsistent / no” (0: both positive and negative results were found or no significant associations were found), or “insufficient” (<3 studies on the specific association were conducted) (37).

In cases where different measures of the same outcome category were assessed within one study (e.g., both stress and sickness absence as indicators of health and well-being within one study), we summarized the findings of the associations among WTC and these different measures into a single rating (see figure 2 for our decision rules regarding the calculation of the single rating). Similarly, when the same association was assessed for different samples within one study, the same decision tree (see figure 2, but with samples rather than measures) was followed to develop one single rating. This single rating can either be positive, negative or zero. The
rating can then be included in the SIC formula to extract the strength of evidence for the association between WTC and this outcome category.

In cases where one association was tested by multiple analyses, we gave priority to the most advanced statistical test (e.g., regression analyses were given priority over correlations and regression models with more control variables were given priority over more simple models) in assessing evidence for associations. As we were interested in main effects of WTC, interaction analyses were not considered.

In one study (38), the authors reported measurement of relevant outcome variables but provided no test statistics for some of the potential associations with WTC. We interpreted the absence of reported results as support for a non-significant finding. In two other studies (39, 40), authors interpreted marginally significant associations (P>0.05) as meaningful effects. In this review, we retained an alpha level of 0.05 for distinguishing significant versus insignificant findings, and marginal effects or associations were therefore considered as null-findings.

Based on these decision rules, the first author rated all reported associations as either positive, negative or zero. Unclear cases were discussed with three other authors.

Categories of worktime control

The 63 papers showed a wide diversity of measures of WTC; we identified five main categories. In the first category, 31 papers assessed WTC in a “global way”, i.e., from the question(s) asked, it could not be determined what specific form of WTC was exercised at the job. Example items included “In general, how much control do you have in deciding when you perform your job?” and “To what extent do you have control over scheduling your working hours?”. In the current review, these studies were categorized as studies of global WTC (15, 25–29, 32, 33, 38, 40–61).

In the second category, 13 papers assessed multiple specific subdimensions of WTC (e.g., control over breaks, starting and ending times, days off, vacation, etc.) but then summed or averaged the scores of these various questions into one overall score of WTC. In this review, these papers comprised the category of studies on “multidimensional WTC” (16, 19, 20, 21, 24, 30, 31, 39, 62–66).

The remaining studies examined the association of a specific subdimension of WTC with relevant outcome variables. In our third category, 13 papers focused on control over daily working hours (22, 23, 67–77). An example question is: “To what extent are you able to influence the length of a work day, and the starting and ending times of a workday?”. We labeled this category “studies on flextime”.

The fourth category of six papers specifically analyzed control over days off or holidays (22, 42, 68, 74, 78, 79). An example question is: “Are you free to decide when to take holidays or a day off?”. We defined this category as “studies on leave control”.

The fifth group of seven papers assessed effects of other subdimensions of WTC (i.e., interruptions during work time for personal matters (27, 76), control over

---

**Table 1.** Strength of evidence for the associations studied in this review, based on the number of studies that assessed each association and the corresponding standardized index of convergence (SIC). [+++/- = strong evidence for a positive/negative association; ++/- = moderately strong evidence for a positive/negative association; +/- = limited evidence for a positive/negative association; 0 = inconsistent evidence or no evidence.]

<table>
<thead>
<tr>
<th>Number of studies</th>
<th>-1.00– -0.60</th>
<th>-0.59– -0.30</th>
<th>-0.29– 0.29</th>
<th>0.30–0.59</th>
<th>0.60–1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength of evidence</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>SIC value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–2</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>3–5</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>≥6</td>
<td></td>
<td></td>
<td></td>
<td>+ + +</td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 2.** Decision rules regarding the calculation of a single rating based on different measures of the same outcome variable within one Study. x represents a positive or negative association; y represents an association that contradicts x; o represents no association. Note. The same decision tree applies in case one study examines one association among different samples (instead of measures).
overtime (12, 66), breaks (80, 81), and a “flexible working hours and compressed working schedule combination” (82)]. This group of papers will be discussed as the “other subdimensions of WTC” category.

Some papers (22, 27, 42, 66, 68, 74, 76) included measurements of more than one subdimension of WTC without converging these into a single “multidimensional score”. Instead, these studies linked the separate subdimensions to outcome variables (eg, paper 68 separately examined the association between flextime and leave control, on the one hand, and outcome variables on the other). These studies were accordingly included in more than one WTC measurement category.

Detailed information on the studies within the various WTC categories and their reported associations with employee outcomes can be found in tables X1–X5 (in the Appendix: http://www.sjweh.fi/data_repository.php).

Categories of outcome variables

The first category of outcome variables was work–non-work balance, including work–home interference, and work–non-work conflict, balance, and enrichment. The second category was health/well-being, comprising indicators of stress, burnout, affective well-being (eg, depression, anxiety), fatigue, sleep, sickness absence, and general health. The third category of outcome variables comprised job-related outcomes: measures of job motivation, satisfaction, performance, and commitment, and actual or intended turnover.

We examined the evidence for relations between WTC and these three outcome categories (work–non-work balance, health/well-being, and job-related outcomes). When the number of studies was too low to allow for any interpretation of evidence (ie, ≤2 studies per outcome category, see table 1), SIC were not reported (notation: insufficient evidence).

For reasons of consistency, we also reported SIC values for the category “other subdimensions of WTC”. It should be noted that these SIC values reflect associations with various subdimensions of WTC and provided little proof of effects of the various individual subdimensions of WTC incorporated within this category.

Results

Descriptive information

Detailed characteristics and findings of the 63 retrieved papers are summarized in tables X1–X5 (http://www.sjweh.fi/data_repository.php). Twenty-seven papers covered heterogeneous working populations comprising mixed age, occupations, and gender. Of the remaining 36 papers, 11 included predominantly (≥80%: 23, 48, 49, 59) or exclusively (15, 38, 39, 42, 46, 50, 64) female samples. Four studies were predominantly (≥80%: 12, 57) or exclusively (40, 66) male samples. Twenty-seven papers addressed samples of specific job categories (16, 32, 33, 38, 40, 43–51, 56–59, 61, 62, 66, 67, 71, 77, 79, 80, 81). Almost all studies were conducted within the US, Europe, or Australia. Studies that cover other countries were scarce (N=4: 52–54, 68).

Of the 63 included papers, 46 used a cross-sectional design (12, 15, 16, 25–31, 41, 43–46, 49, 51–63, 65–77, 79–82) and 11 papers (representing 6 studies) employed a longitudinal design (19–24, 39, 42, 47, 64, 78). Four of these (39, 47, 64, 78) reported both cross-sectional and longitudinal data and, for these studies, we will consider both cross-sectional and longitudinal associations.

Finally, 7 papers were found that addressed 5 intervention studies (32/33, 38, 40, 46/50, 48). These five intervention studies included pre- and post intervention measurements among an intervention group as well as a control group. They did not use randomization to allocate participants to control and intervention conditions.

All studies used self-report methods for data collection. In addition to individual self-report measures, one paper also assessed WTC on the work unit level (24). Furthermore, papers 19–24 (all based on the 10-Town Study) also included registered data for assessing certain health variables (sickness absence and disability pension), while one study (81) conducted physiotherapeutic examinations to assess musculoskeletal disorders. Two intervention studies supported self-reported health indicators with physiological measures [ie, of heart rate and blood pressure (40) or biomarkers collected from blood samples (46)]. To assess job-related outcomes, one paper reported registered data on turnover (33), while another used both self ratings and peer-ratings to measure employee performance (65).

Evidence for associations between each worktime control and outcome categories

Table 2 summarizes the findings of this review.

Global worktime control

Work–non-work balance. Of the 16 cross-sectional studies on global WTC and work–non-work balance, 11 reported that global WTC is positively associated with work–non-work balance (15, 25/26, 27/28/29, 41, 42, 44, 47, 51, 54, 58, 61), whereas one study reported mixed findings among different samples (53), and four studies reported no relation (43, 45, 56, 60). These results yielded a SIC (N=16) of 0.69, providing strong evidence for a positive cross-sectional relation between global WTC and work–non-work balance. Two longitudinal studies (42 and 47,
the latter of which included one study that additionally employed a cross-sectional design) yielded no temporal relations between global WTC and work–non-work balance (insufficient evidence).

One intervention study (48) among mostly female nurses, showed that work–non-work balance improved after the introduction of higher levels of WTC (ie, self-scheduling). Another intervention study (32) also showed an improved work–non-work balance after the introduction of increased WTC among white-collar workers. A third intervention study (46: self-scheduling among female eldercare workers) reported no effects of increased WTC on work–non-work balance. Together, these intervention studies yielded a SIC (N=3) of 0.67, providing moderately strong evidence that increased WTC results in improved work–non-work balance.

**Health/well-being.** Of 11 cross-sectional studies on global WTC and health and well-being, 4 (25, 29, 51, 57) reported positive associations (25: stress and burnout, affective well-being, and general health; 29: affective well-being; 51: stress; 57: musculoskeletal symptoms). Six (27, 43, 49, 52, 60, 61) reported no association with indicators of health and well-being (27, 49, 52: general health; 43: fatigue and general health measures; 60: sickness absence; 61: affective well-being, fatigue and general health). One study (53) reported an unfavorable association among employees from Singapore, but no association among US employees (ie, with affective well-being). Together, these studies yielded a SIC (N=11) of 0.27, providing inconsistent evidence for a positive association between global WTC and favorable health and well-being indicators.

Two intervention studies among mostly (48) or exclusively (46/50) female workers in the healthcare sector showed no effects of WTC improvements on health and well-being (46: stress; 48: vitality, stress and general health; 50: sleep quality). Two intervention studies (38, 40) reported significant effects of increased control over working schedules for some indicators of well-being and health, but not others [38: (midwives) decreased tiredness during night shifts but no significant changes in mental stress and mental strain; 40: (male airline maintenance personnel) significant decrease in blood pressure but no effects on heart rate, sleep-outcomes or general health]. Together, these intervention studies yielded a SIC (N=4) of 0.00, providing no clear evidence for effects of increased WTC on health and well-being outcomes.

Two longitudinal studies (33, among white-collar workers; 48, among predominantly female nurses) showed a positive impact of increased WTC on job-related outcomes (33: both registered and intended turnover; 48: job satisfaction) (insufficient evidence). A longitudinal study (47) reported no association (ie, with work engagement; insufficient evidence).

### Multidimensional worktime control

**Work–non-work balance.** Of three cross-sectional studies, two (62, 65) reported a positive association between higher multidimensional WTC and work–non-work balance, whereas one (64) failed to find an association. Together these studies yielded a SIC (N=3) of 0.67, providing moderately strong evidence for a positive association between multidimensional WTC and work–non-work balance.

**Health/well-being.** Of four cross-sectional studies, two (63, 65) reported a positive association between multidimensional WTC and health and well-being (63: general health; 65: recovery-related well-being) whereas the other two (30/31, 66) reported no association (30/31: stress; 66: affective-well-being, fatigue and general health). These studies yielded a SIC (N=4) of 0.50, thus providing limited evidence for a positive association between multidimensional WTC and health and well-being.

One longitudinal study (19/20/21/24) reported positive associations between multidimensional WTC and various health and well-being indicators (ie, with sickness absence, general health, affective-well-being and risk of disability pension) whereas a second longitudinal study (64) reported no association (ie, with general health; insufficient evidence).

**Job-related outcomes.** One cross-sectional study (16) reported mixed findings regarding the association between multidimensional WTC and job-related outcomes (ie, positive association with affective commitment, no association with job satisfaction). Two cross-sectional studies (39, 65) failed to find any significant association (39: job satisfaction; 65: self- and peer-assessed job performance). Together these studies provided a SIC (N=3) of 0.33, providing limited evidence for a positive association between multidimensional WTC and job-related outcomes.

Two longitudinal studies (39, 64) found no association with job-related outcomes (39: job satisfaction; 64: voluntary turnover; insufficient evidence).
Review on worktime control

Flextime

Work–non-work balance Of the eight cross-sectional studies that examined the association between flextime and work–non-work balance, three (67, 69, 73) reported a significant positive association between flextime and work–non-work balance, whereas five studies (71, 72, 74–76) reported no significant association. Together, these studies yielded a SIC (N=8) of 0.38 and provided moderately strong evidence for a positive association between flextime and work–non-work balance.

Health/well-being. From three cross-sectional studies on flextime and health and well-being, two (68, 70) reported a positive association with health and well-being (68: affective well-being, sleep, and recovery-related outcomes; 70: stress symptoms). A third cross-sectional study (71) reported no association (ie, with affective and physical well-being). Together, these studies provided moderately strong evidence for a positive association between flextime and health/well-being [SIC (N=3) 0.67].

One longitudinal study (22/23) showed an association between flextime at baseline and health/well-being at a 3–4-year follow up (ie, lower sickness absence; insufficient evidence).

Job-related outcomes. Of four cross-sectional studies, two (76, 77) reported a positive association between flextime and job-related outcomes (76: job satisfaction; 77: job satisfaction and organizational commitment) while a third cross-sectional study (73) reported mixed associations with job-related outcomes (ie, a positive association with job satisfaction, no association with self-rated job performance). One study (67) reported no association (ie, with affective commitment). Together, these studies yielded a SIC (N=4) of 0.75, providing moderately strong evidence for a positive association between flextime and job-related outcomes.

Leave control

Work–non-work balance. Two cross-sectional studies (74, 78) reported a positive association between leave control and work–non-work balance (insufficient evidence). One of these studies additionally reported a...
positive longitudinal association between leave control and work–non-work balance (78). A second longitudinal study (42) reported a favorable association as well (insufficient evidence).

**Health/well-being.** Of two cross-sectional studies, one (68) reported a positive association with health and well-being (ie, affective well-being, sleep and recovery-related outcomes), whereas the second (79) reported no association (ie, with affective well-being, sleep-quality, recovery, and general health; insufficient evidence). One longitudinal study (22) reported a positive association between leave control and health and well-being (ie, sickness absence; insufficient evidence).

**Job-related outcomes.** No studies were found that assessed leave control in relation to job-related outcomes.

Other subdimensions of WTC

**Work–non-work balance.** One cross-sectional study (82) reported a positive link between access to flextime or compressed workweek schedule and work–non-work balance. Two other cross-sectional studies (27, 76) reported a positive association between control over interruptions for personal matters during working hours and work–non-work balance. Together, these studies yielded a SIC (N=3) of 1.0, providing moderately strong evidence for a positive association between other subdimensions of WTC and work–non-work balance. However, this result is difficult to interpret, as it does not provide insight into the associations between the individual types of WTC and work–non-work balance.

**Health/well-being.** Of two cross-sectional studies on control over overtime (12, 66), one reported a favorable association with health and well-being (12: lower fatigue), whereas the other found significant associations for some outcomes but not others (66: a positive association with general health, no association with affective well-being or fatigue). Regarding control over breaks, one cross-sectional study found a favorable association with health and well-being (81: lower musculoskeletal symptoms as assessed by both self-report and medical examination), whereas another cross-sectional study found mixed results (80: higher general well-being, no association with musculoskeletal symptoms). For possibilities to interrupt work for personal matters during work hours (27), no overall association was found (ie, with general well-being). Together, the studies resulted in a SIC (N=5) of 0.60, representing moderately strong evidence for a positive association between other subdimensions of WTC and health and well-being indicators. Again, this result is hard to interpret, as it does not inform us about associations between health and well-being and specific types of WTC.

**Job-related outcomes.** One cross-sectional study (82) reported a positive association between access to flextime or compressed workweek schedule and favorable job-related outcomes (ie, job satisfaction and turnover intention). In addition, favorable job-related outcomes were associated with control over: (i) overtime (12: job satisfaction), (ii) breaks (80: job satisfaction), and (iii) interruptions for personal matters during working hours (76: job satisfaction). These studies together provided moderately strong evidence for a favorable association between other subdimensions of WTC and job-related outcomes [SIC (N=4) 1.0]. Once more, this result is hard to interpret as it covers various subdimensions of WTC.

**Discussion**

This systematic review examined current empirical evidence regarding the association between WTC and (i) work–non-work balance; (ii) health/well-being; and (iii) job-related outcomes. A total of 53 studies, published 1995–2011, were included in the review.

**Research question 1**

*How strong is the empirical evidence regarding the association between (categories of) WTC and indicators of work–non-work balance, health/well-being, and job-related outcomes?* The strongest and most consistent evidence (ranging from moderately strong to strong) was found for a positive association between WTC and work–non-work balance. The association was found for both global and multi-dimensional measures of WTC, as well as for flextime. It was observed in both cross-sectional and intervention designs.

Regarding cross-sectional studies on indices of health/well-being, the review showed inconsistent evidence for a positive association with global measures of WTC, limited evidence for positive associations with multidimensional WTC, and moderately strong evidence of positive associations with flextime. The limited number of intervention studies showed no evidence for overall effects of WTC on health/well-being (although it should be noted that significant effects were found for particular individual indicators of health/well-being).

In the analysis of job-related outcomes, results ranged from limited evidence of a positive association with multi-dimensional measures of WTC, up to moderately strong evidence of positive associations with global WTC and flextime. Two intervention studies...
(33, 48) observed positive effects of increased WTC on job-related outcomes, but the number of intervention studies was too low to allow firm conclusions on the effectiveness of WTC on job-related outcomes.

Our analysis of specific subdimensions of WTC identified moderately strong evidence for positive cross-sectional associations between flextime and work–non-work balance, health/well-being, as well as job-related outcomes. Promising results were also found for other specific subdimensions of WTC (eg, leave control, control over breaks, and control over overtime). However, the number of studies on any of these specific subdimensions of WTC was too low to provide sufficient evidence for a positive association.

Research question 2

In case of significant associations between (categories of) WTC and these indicators, how strong is the empirical evidence that these associations are causal in nature? In order to draw inferences regarding causality, four methodological requirements have to be met: (i) significant associations; (ii) temporal ordering; (iii) theoretical plausibility for the presumed causal relationships; and (iv) exclusion of rival hypotheses (83). Only longitudinal and intervention studies meet the second requirement, with the latter providing the strongest opportunity to assess causal associations. Within the current review, three intervention studies (32, 46, 48) focused on the association between global measures of WTC and work–non-work balance, of which two identified a significant positive association, thereby fulfilling requirements (i) and (ii). The significant associations found are consistent with the time- and recovery-regulation mechanisms, and the occupational health theories that were outlined in the introduction. Hence requirements (iii) and (iv) have been fulfilled. Therefore, we conclude there is moderately strong evidence that higher global WTC [ie, a general increase of WTC (32) or introduction of self-scheduling (46, 48)] causes an improvement in work–non-work balance.

Four intervention studies (38, 40, 46/50, 48) focused on the effects of global WTC on several indices of health/well-being. As discussed above, these studies found no evidence for overall effects. However, it must be noted that some specific individual indicators of health/well-being [ie, tiredness during the nightshift (38), blood pressure (40)] did positively change as a result of WTC interventions. We therefore tentatively conclude that WTC may have positive causal effects on health/well-being, but more intervention research is needed to examine which specific health/well-being indices are sensitive to changes in WTC.

For all the other associations studied, it was not possible to draw causal inferences, due to insufficient evidence (ie, a scarcity of intervention studies) or sub-optimal research design quality.

Strengths and limitations of the review

One of the keys strengths of the current review was its breadth of focus. Firstly, it examined a wide range of operationalizations of WTC, while excluding related but separate constructs such as spatial control (ie, control over where to work). Secondly, it covered a relatively broad range of study designs, including not only longitudinal and intervention studies, but also cross-sectional designs. While acknowledging the restrictions of cross-sectional designs regarding causal inferences, we would argue that such studies provide valuable information regarding possible effects of WTC. Thirdly, the review examined a broad range of theoretically and practically relevant outcome variables. Finally, the review also provided a detailed picture of the associations between specific subdimensions of WTC and several outcomes.

A number of limitations should also be noted. Firstly, some of the observed associations were based on the same dataset and were, therefore, not independent. For example, some studies identified both cross-sectional and longitudinal associations between WTC and a specific outcome within the same sample (eg, 39, 47), and hence the two findings were not independent. Moreover, in some studies (eg, 27, 65, 71), the observed associations between WTC and several outcome categories (eg, work–non-work balance and health/well-being) were not independent as they were based upon the same sample of respondents. This overlap in data could, in theory, result in an overestimation of the favorable associations with WTC. However, when looking at table 2, it can be seen that several studies with multiple outcome categories showed contrasting findings for work–non-work balance, health/well-being, and/or job-related outcomes. For instance, one study (43) showed positive findings regarding work–non-work balance and non-significant findings regarding health/well-being. Similar variation in findings over outcome categories was observed within several other studies (eg, 27, 47, 52, 61, 65). Moreover, the same picture emerges when looking at studies that examined both cross-sectional and longitudinal associations within the same dataset. For instance, two studies (42, 47) found significant positive cross-sectional associations between global WTC and work–non-work balance, but longitudinal associations between these variables within the same dataset were non-significant. Thus it seems unlikely that overlapping data has resulted in a marked overestimation of favorable associations.

A second limitation was that the reviewing process did not take into account the quality of the measurements used in the studies, as to do so would have overly complicated the analysis. Yet, it must be noted...
that several WTC studies (eg, 25/26, 42) included sub-optimal one-item or non-validated measurements of the central research constructs (WTC and outcome measures). A multi-dimensional measurement of WTC may be regarded as more valid than a single-item global measure and the same is true of the various outcomes examined. Such crude measures of global WTC are incapable of capturing the complexity of the work situation of employees who, for example, might have high control over certain aspects of their working time (eg, when to take days off) but no control over their daily work hours (eg, taking breaks, starting and finish times, overtime). It is also notable that several studies (30/31, 63) were based on the European Union surveys of working conditions which rely on measures that have not been psychometrically validated. Such data also suffer from a number of other potential problems (eg, cultural differences, issues of translation, labor market differences) which may introduce biases into the findings.

Thirdly, it remains unclear from the review results whether the associations between WTC and the outcomes were independent of, or mediated by, other psychosocial work characteristics. For example, few of the studies took into account (ie, adjusted for) associations with general job control. Thus we cannot be certain that the observed associations would have held if the effects of WTC had been isolated in these studies, nor can we infer whether WTC directly influenced outcomes or that the effects were mediated by a change in the (psychosocial) work environment.

The fourth and final limitation concerns publication bias in favor of statistically significant findings, which may be more easily published than null findings. Hence the ratio of positive associations to null effects may be somewhat inflated in the current review. However, it is notable that only 1 of the 53 studies found significant negative (unfavorable) associations between WTC and the outcomes examined. Thus it may be concluded that high WTC is generally not related to adverse employee outcomes.

Recommendations for future research: a research agenda

On the basis of our findings, we consider the topic of WTC to be fruitful ground for at least another decade of research. More research is needed to draw definite conclusions about the causal influence of WTC on relevant outcomes (ie, work–non-work balance, well-being/health, and job-related outcomes). We propose three recommendations that may guide future research on this topic.

Firstly, we recommend that researchers not only examine the effects of general (global/multidimensional) WTC, but also examine the effects of the specific sub-dimensions of WTC (flextime, leave control, break control, control over overtime), to determine which subdimensions are most strongly related to which types of outcome variable. It can be hypothesized that some subdimensions mainly work at the level of the recovery-regulation mechanism (eg, control over breaks) and will therefore be chiefly related to recovery-related outcomes (eg, fatigue and vitality). Other subdimensions may function at the level of the time-regulation mechanism (eg, flextime) and may therefore primarily affect indices of work–non-work balance. Finally, some subdimensions may affect outcomes through both regulation mechanisms (eg, leave control and control over overtime) and may therefore affect a broader spectrum of outcome variables. A broad measurement of all subdimensions of WTC is also relevant since the effects of a single subdimension (eg, leave control) may to some extent also depend on the level of WTC over the other subdimensions. That is, high leave control may not promote a favorable work–non-work balance or better health if control over other subdimensions of WTC (eg, control over start and finish times, or overtime) remains limited. As such, a complete measurement of all relevant WTC aspects is recommended if one wishes to establish a valid insight into separate and combined effects of subdimensions of WTC.

Secondly, the WTC literature would benefit from studies with high quality designs, especially longitudinal and intervention studies, as these designs allow causal inferences to be made. Regarding the focus of future intervention studies, we recommend more attention to be paid to modern WTC practices that are currently popular within organizations, eg, self-scheduling (also known as self-rostering) and boundaryless work. Self-scheduling is mostly applied in shift work settings and may provide shift workers with more freedom regarding their work schedule. Boundaryless work has recently become popular in office settings among white-collar workers. It includes a combination of extensive WTC and spatial flexibility, with employees being able to decide when and where to work. As noted in the review, self-scheduling shows some promising results for work–non-work balance and job satisfaction (48), although there were null-findings for indices of recovery and general health (38, 40, 46/50, 48).

Thirdly, many studies relied primarily on self-reports to collect data on WTC and several outcome measures. Spector (84) has shown that reliance on self-reports does not necessarily result in problems of common method bias. Nevertheless, in future studies, it is desirable that researchers also include other data sources (eg, administrative data on sickness absence, cf. 19–24) and measures of WTC using multiple assessors (eg, assessment by the employee and the supervisor; cf. 24).

A final recommendation is to distinguish between
relevant subgroups when examining the effects of WTC. Although WTC is assumed to be universally beneficial, it is likely that employees with greater family responsibilities (eg, women) and those with greater need for recovery (eg, older workers) will gain the most. Women still tend to be primarily responsible for home and family obligations. One study by Ala-Mursula and colleagues (23) indicated that providing women with greater levels of WTC helped them maintain favorable health and well-being, even when working relatively long hours. However, it is also worth noting that flexibility could have negative consequences for women. For example, women may end up engaging in more non-work responsibilities, rather than using the increased time control to recover more completely and lower stress and strain outcomes (85).

Many governments are seeking to increase and extend the labor participation of older workers (86). WTC may help keep older employees actively involved in the work community while meeting their personal needs for more free time and time for recovery. Older employees may be more willing and able to remain working if they can decide the quantity and distribution of their work hours. Only one of the papers included in the current review examined age as a possible moderator (30) in the association between WTC and outcomes, and none of the studies specifically focused on older employees.

In conclusion, this review has shown that there are theoretical and empirical reasons to view WTC as a promising tool for the maintenance of employees’ work–non-work balance, health and well-being, and job-related outcomes. At the same time, however, the current state of evidence allows only very limited causal inferences to be made regarding the impact of enhanced WTC.

Acknowledgments

The authors would like to thank ZonMW for funding.

References


http://dx.doi.org/10.1007/s00420-009-0411-z.


http://dx.doi.org/10.1080/02678370801984927.


Received for publication: 21 December 2011