Advances in Accounting 61 (2023) 100646

Contents lists available at ScienceDirect

Advances in Accounting

journal homepage: www.elsevier.com/locate/adiac

Work-life balance in public accounting: An experimental inquiry into supervisor support for subordinate career progression

Mary Sasmaz, Timothy J. Fogarty *

Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106, United States of America

ARTICLE INFO

Accepted by Dennis Caplan

Keywords:
Work-life balance
WLB
Non-financial metric
Balanced scorecard
Gender
Audit firms

ABSTRACT

Work-Life Balance (WLB) continues to be a concern of audit professionals because the long work-hours environment can have negative effects for both individuals and organizations. Audit firms have continuously committed to helping employees with the creation of work-life balance and well-being programs. The purpose of this study is to determine whether the firm's official commitment to work-life balance is reflected in supervisors' evaluation of subordinates. This study conducts a between subjects experiment using actual audit supervisors as participants to capture responses to ways that a hypothetical staff person might pursue WLB. As part of this, a hypothetical non-financial WLB metric used as part of the formal performance evaluation process is examined as a potential tool for strengthening the effectiveness of audit firm investments in WLB. The results show that WLB alternatives still have negative career consequences, and these consequences would not be mitigated by the use of a formal WLB performance evaluation metric. Although career consequences of WLB are not significantly related to gender, performance evaluation is not gender neutral.

1. Introduction

Work-Life Balance (WLB) has been a concern in professional service careers for many years. This study addresses three main issues important to understanding the impact of WLB in the accounting profession. First is whether the specific format of WLB utilized by a staff member has a differential impact on supervisors' career progression recommendations for that person. Second, we explore whether the historic gender-specific implications of WLB on career progression continue to characterize the profession. Third, we explore the potential of a formal WLB metric included in performance appraisal to mitigate the detrimental career effect of WLB choices.

Work-life balance has been a concern of audit professionals for years, given the long work-hours environment associated with the occupation (Anderson-Gough, Grey, & Robson, 2000; Ladva & Andrew, 2014). Long work-hours reduce work-life balance and can have negative effects for both individuals and organizations. Consequences include stress (Cooper, Davidson, & Robinson, 1982), fatigue (Ono, Watanabe, Kaneko, Matsumoto, & Miyao, 1991), and other health issues including increased coronary heart disease risk (Virtanen et al., 2012). Long hours also have negative job performance consequences such as reduced productivity (Pencavel, 2015). WLB issues have been suggested as a potential significant contributor to retention problems within accounting firms (American Institute of Certified Public Accountants (AICPA), 2004, PricewaterhouseCoopers (PwC), 2013, Barry, 2015, Buchheit, Dalton, Harp, & Hollingsworth, 2016, Mendlowitz, 2018). The negative impact of excessive turnover on audit quality suggested by the PCAOB (2013) highlights another practical concern of these firms. Furthermore, recent surveys of practicing accountants suggest WLB considerations may be an under-recognized determinant of both entry into and turnover from public accounting (American Institute of Certified Public Accountants (AICPA), 2004, Barry, 2015, PricewaterhouseCoopers (PwC), 2013). Recent generational transition observations heighten these concerns (Adamson, 2013; Boomer, 2018).

Audit firms have committed to helping employees find desired WLB with the creation of work-life balance and well-being programs. Such programs are intended to encourage employee recovery from the difficulty of work, which has been linked to increased job performance (Binnewies, Sonnentag, & Mojza, 2010) and improved health outcomes. Improved WLB is believed to improve staff retention within audit firms (Tysiac, 2017). With WLB specifically believed important to the Millennial generation (PricewaterhouseCoopers (PwC), 2013, Twenge, 2010), a group now representing a significant portion of the firm's workforce, the effectiveness of WLB and well-being programs is

* Corresponding author.
E-mail addresses: mbs26@case.edu (M. Sasmaz), tjf@case.edu (T.J. Fogarty).

https://doi.org/10.1016/j.adiac.2023.100646
Received 3 July 2021; Received in revised form 4 February 2023; Accepted 18 February 2023
Available online 15 March 2023

especially important to audit firms (see also Lightbody, 2009). However, the client-centered business model and management goals (efficiency, profitability, growth) of public accounting firms may encourage personnel behavior that contradicts WLB efforts (Ladva & Andrew, 2014). Because WLB disproportionately concerns female employees, its success will also impact accounting firms' ability to capitalize on recent gender diversity gains in hiring staff.

One purpose of this study is to determine whether the commitment to work-life balance that firms have made through investments in employee programs is reflected in supervisors' evaluation of subordinates for purposes of career progression. Without the support of WLB by immediate supervisors, efforts at the firm's management level to encourage balance will be ineffective. In this way, this research seeks to inform what many see as a puzzle. Public accounting firms are often named as "Best Places to Work" by mass media (e.g., Fortune, 2021). However, anecdotal accounts of those who depart from employment at these firms often cite unacceptable WLB as a major reason for leaving.

Historically, WLB has been viewed as a female-centric issue given traditional dominant roles of women in household and childrearing activities. With the broadening involvement of men in the home and with childcare, and shifting generational desires regarding personal time use, WLB has become an issue for both genders in the professional work environment. Recent surveys suggest that WLB is an important concern for both males and females (PricewaterhouseCoopers (PwC), 2013). However, the equality between genders regarding WLB is an empirical question.

Managing a professional service firm requires the serious use of time budgets and management of staff to meet deadlines. These pressures compete against WLB efforts. Immediate supervisors are tasked with delivering audit work on time and within budget in an environment where upper management has publicly supported WLB. This study explores whether a non-financial WLB metric included in the formal performance evaluation process can counteract the competing pressures of budget and deadlines to support the effectiveness of audit firm investments in work-life balance and well-being programs. Specifically, the existence of a balanced-scorecard-type dimension is experimentally elevated from a vague firm-level concern to one more relevant to front-line supervisors.

The results, produced by the administration of experimental materials to practicing auditors in large firms in the U.S., suggest that despite years of significant investment in work-life balance programs by accounting firms, WLB utilization continues to likely result in slower career progression for its users, as such is influenced by supervisor evaluations. However, the magnitude of this effect depends on the specific type of WLB that is used. The career consequences associated with the use of WLB is not significantly related to staff member gender, though differences exist for the specific career outcome in question. The use of a balanced-scorecard-type evaluation tool does not alter the career progress results associated with WLB choices. However, some evidence exists that the use of that tool is not gender neutral

The remainder of this paper is organized onto four sections. Section II provides a brief literature review of the areas that past work has informed. This supports the formulation of new hypotheses that represent the contribution of this paper. Section III describes the study conducted to test the hypotheses. This part of the paper includes a description of the measures used for the collection of the data. Section IV summarizes the findings. The paper concludes in Section V with a discussion of the results, their implications, their limitations, and future research opportunities.

2. Literature review and hypothesis development

The WLB construct has evolved over time from a work-life conflict concern to a work-life harmony perspective (Jain & Nair, 2013). The progression of the WLB construct is consistent with the overall evolution of the business and economic environment. Globalization and technological advances have resulted in the "dilation of the boundary between work and family" especially in the services sector (Jain & Nair, 2013). Specifically, with respect to the accounting profession, the Sarbanes-Oxley Act of 2002 significantly expanded the responsibility of auditors (Cohen, Krishnamoorthy, & Wright, 2010). The increase in workload results in strained WLB for auditors (Buchheit et al., 2016), already stressed by the seasonality of audit work (Sweeney & Summers, 2002). Consequently, a large segment of the accounting profession routinely suffers from inadequate WLB (Gullapalli, 2005). Achieving WLB can no longer just be a personal goal of employees, but also requires intentional involvement from employers. Rather than a problem faced by selected employees, WLB has become a generalized objective for all.

In response to WLB concerns, many accounting firms including all of the Big Four, announced WLB and well-being programs to help employees achieve WLB. Such programs are intended to encourage employee recovery. Recovery has been linked to increased job performance (Binnewies et al., 2010). Recovery activities increase an individual's WLB because people are aware that intense work periods will be quickly compensated with time available to pursue non-work objectives. One example of a Work-Life Balance Program within accounting firms is an alternative work arrangement (AWA). AWAs consist of flexible schedules, modified work weeks, reduced hours schedules, as well as telecommuting opportunities. Well-being programs offer rewards to employees who participate in healthy behaviors such as getting medical check-ups and exercising regularly. For example, Grant Thornton announced a Work-Life Balance Program in 2015, promoting flex-time for all employees (Grant Thornton, 2015). By the end of 2016, 96% of large national accounting firms (firms with revenues greater than $10 million) and at least two-thirds of other accounting firms with revenue above $200,000 were offering such flexibility (Tysiac, 2017). PricewaterhouseCoopers' current website boasts quality of life provided for employees through both informal everyday flexibility programs (such as year-round flex days and teaming culture) as well as formal flextime programs such as reduced hours and compressed workweeks (PwC, 2020). A 2018 study indicates that 58% of accounting firms nationwide (up 19% from the prior year) have implemented "Optional Saturdays" rather than required weekend work. Firms indicate that actively managing a group of WLB programs is now a priority (American Institute of Certified Public Accountants (AICPA), 2019a).

Less certainty exists about whether the commitment to work-life balance that firms have made through investments in employee programs is reflected proportionately in supervisors' evaluation of subordinates for purposes of career progression. Without the support of WLB by supervisors, efforts at the firm's management level to encourage balance will be less effective. The implementation of work-life balance and well-being programs is a significant change from the traditional work policies and perceptions that previously defined accounting firm culture. Immediate supervisors are tasked with delivering audit work on time and under budget. This demonstrates a potential conflict between short-run tangible objectives and the more long-term intangible goals of official firm programs promoting WLB.

2.1. WLB approach and career consequences

2.1.1. Alternative work arrangements

Despite association with reduced turnover (Almer & Kaplan, 2002), past research has shown that participation in WLB programs such as the alternative work arrangement (AWA) have been perceived as detrimental to career success within public accounting firms (Cohen & Single, 2001; Johnson, Jordan Lowe, & Reckers, 2008). This suggests that despite publicly announced organizational support for such programs, the traditional work schedule is still considered by supervisors to be more deserving of rapid career progression for staff. Therefore, current participation in an AWA program should be expected to hinder recommendations for promotions, bonuses, and salary increases.

On the other hand, evolutionary theorists would argue that organizations change due to environmental demands for survival (Kezar, 2001). Globalization, technological advances, and differences in perceptions of employees due to generational shifts may have led to an environment more accepting of WLB program participation. This might make it less likely that negative consequences for career success would be attached. Buchheit et al. (2016), in a survey of 1063 CPAs, find that 33% to 64% of big firm participants perceived strong organizational support for AWAs, which included several specific types of arrangements (e.g., flextime, part-time, telecommuting). However, whether organizational support for such programs is evident in current immediate supervisor perceptions, and the resultant career progression decisions regarding subordinates, remains an open question.

While promotions in accounting firms are decisions typically made during a meeting of managers or partners, a prominent basis for their decisions is the information documented in performance appraisals on various projects. Since supervisors prepare performance appraisals, perceptions of supervisors become an important factor in the promotion and annual review processes. Manager perceptions have been found to directly impact performance evaluations (Pelletier & Vallerand, 1996). Furthermore, within the accounting field, negative supervisor perceptions of WLB programs have been linked to adverse career consequences specifically related to scheduling and potentially reduced assignment of challenging work (Johnson et al., 2008).

Results from the Johnson et al. (2008) study include some ambiguity in how participation in an AWA may impact formal performance feedback. There, managers were provided performance information about a subordinate, and notice about whether the subordinate did or did not participate in an AWA. Informally, the manager indicated the likelihood they would schedule the subordinate on a future job, as well as their perception of how challenging the future work assignments that this subordinate should receive. The manager also was asked to provide a formal performance evaluation ranking for the specific assignment. These authors did not find statistically significant differences in the rankings between subordinates who participated in AWAs and those who did not for the formal job performance evaluations. However, the authors did find statistically significant differences in the informal feedback results regarding future job scheduling and the future distribution of challenging work. This suggested that AWA participation was perceived negatively by supervisors and was likely to have negative career consequences. The authors suggest that the lack of statistical significance of AWA use on formal performance appraisals may be due to managers' desire to be "politically correct" by harmonizing with the firm's WLB initiatives. The lack of significance could also be evidence of evolutionary change within the firms as it adjusts to the WLB expectation of the current workforce.

2.1.2. Weekend recovery

Our appreciation for WLB in public accounting should not be limited to AWAs. Another approach taken by some employees to achieve work-life balance is through weekend recovery. Weekend recovery means participating in activities during the weekend that lead an individual to feel physically and mentally refreshed at the end of the weekend (Binnewies et al., 2010). Recovery activities include relaxation as well as detachment such as taking electronic device "holidays" and spending rejuvenating time outdoors. Weekend recovery has been positively associated with job performance after the recovery weekend (Ragsdale & Beehr, 2016). Specifically, weekend recovery has been linked to improved weekly task performance, enhanced personal initiative and more frequent organizational citizenship behaviors (Binnewies et al., 2010), and subsequent higher engagement and burnout resistance (Ragsdale & Beehr, 2016). Weekend recovery possesses relevance to WLB in accounting firms in that it can be unilaterally initiated. It is less extreme than an AWA since it does not have to be negotiated with the firm. As such. Weekend recovery is neither officially approved nor discouraged by firms. Although accounting firms do not take a formal stance regarding a weekend recovery approach to balancing work and life, the setting of boundaries, purposeful time management, and technology distancing suggested in practice journals to professionals looking to achieve balance, would accomplish the desired WLB objectives many have (Davidson, 2018; Little, 2018). Recently, the *Journal of Accountancy*, a publication of the AICPA commonly read by practitioners, included an article encouraging managers to lead by example in achieving WLB and highlighted a manager who removed work email from her cellphone and avoids turning on her computer on the weekend (Hart, 2021). Furthermore, the "flexibility" or "flextime all the time" that firms now boast would lead one to believe that weekend recovery is indeed one approach an accounting professional might implement in the quest for better balance between work and life. The weekend recovery approach to WLB has not yet been formally evaluated in the accounting literature.

Despite the benefits of weekend recovery, public accounting firms have been found to sustain a male stereotype with respect to performance appraisals and career progression that may impede its use (Anderson, Johnson, & Reckers, 1994; Fogarty, Parker, & Robinson, 1998). Recent research and partner statistics suggest the male stereotype remains prevalent in current views of career progression within public accounting firms (Johnson et al., 2008; American Institute of Certified Public Accountants (AICPA), 2015). Research by Williams (2010) and Williams, Blair-Loy, and Berdahl (2013) attribute this to the "ideal worker" norm that is present in the current marketplace. The ideal worker norm suggests that a person be fully committed to the employer without family constraints (Williams, 2000), and therefore be willing to use weekends for work on a regular basis.

2.1.3. The consequences of different types of WLB (H1)

With technology and the effects of globalization having blurred the line between work and personal life, the ideal worker has become viewed as one who is always available (Schulte, 2015). Participation in WLB programs, whether formal firm programs or self-designed informal programs, mitigates the level of one's availability for work-related matters. As one's availability is reduced, adverse career progression consequences may occur. This might result because supervisors will make less positive career progression recommendations for individuals who use weekend recovery to achieve WLB than individuals who do not make a specific effort to achieve WLB. Similarly, individuals who participate in AWAs would be expected to have even lower career progression recommendations since there would be more formal and comprehensive limits to their availability. Furthermore, past studies suggest that formal WLB programs are less acceptable for those in lower ranks in the firms (Johnson, Lowe, & Reckers, 2012; Kornberger, Carter, & Ross-Smith, 2010). In order to discover if sufficient evolution within firm cultures has occurred such that striving to achieve WLB will not result in a negative influence on manager decisions about subordinate career progression, one needs to study both official and unofficial methods of WLB throughout the career of a staff accountant.

One way to remove some of the interpretive ambiguity of the literature is to clarify and expand the decision outcomes solicited from respondents. Recommendation for promotion is a more significant decision than the performance appraisal from a single engagement which was measured in the Johnson et al. (2008) study. Promotion asks respondents for a more holistic judgment about the individual. Another measure of career progression is monetary in nature. If participants were asked to provide their recommendations for the compensation

incentives that accompany performance, they might make such allocations differently than they do with the dispensation of words of praise. Money, unlike praise, is a scarce resource that speaks loudly to its possessors. Ceteris paribus, subordinates who strive to achieve WLB by participating in an AWA program are expected to be recommended less often for promotion than individuals who achieve WLB through weekend recovery or individuals who make no attempt to achieve WLB. Subordinates who use AWAs to achieve WLB are also expected to receive lower amount recommendations of bonuses, and receive lower percentage recommendations for salary increases, because they have explicitly negotiated for larger blocks of non-work time.

The first hypothesis addresses the expected career repercussions of striving for work-life balance in a profession that has historically and traditionally expected a highly committed employee to be always available. For simplicity in interpretation, each WLB approach is considered separately and compared to the individual who makes no effort to achieve WLB. Finally, the two WLB conditions are compared to one another to evaluate the expected ordering effect. Formally put, the first hypothesis in three parts can be stated as follows:

H1a. Individuals who use AWAs to achieve WLB will receive lower career progression recommendations than those who do nothing to achieve WLB.

H1b. Individuals who use weekend recovery to achieve WLB will receive lower career progression recommendations than those who do nothing to achieve WLB.

H1c. Individuals who use AWAs to achieve WLB will receive lower career progression recommendations than those who use weekend recovery to achieve WLB.

2.2. Gender differences on WLB consequences (H2)

The historically prevalent ideal worker norm suggests the ideal worker to be a person fully committed to the employer notwithstanding family constraints (Williams, 2000). Since women are more likely to take on the caregiver role, they have commonly been excluded from being viewed as the "ideal worker" (Williams, 2000; Williams et al., 2013). Although the roles of men and women have been changing within the accounting workforce with respect to childcare responsibilities (PricewaterhouseCoopers (PwC), 2013, American Institute of Certified Public Accountants (AICPA), 2011), the "ideal worker" norm is suggested by some to still typify public accounting (Ely & Padavic, 2020; Williams, 2010).

Men who step out of this societal norm experience professional consequences including lower raises, fewer promotions, and increased risk of termination or of being down-sized (Rudman & Mescher, 2013). Although millennial families tend to be comprised of more dual income earners, which might result in both parents being involved in child rearing at greater levels than previous generations (American Institute of Certified Public Accountants (AICPA), 2011), males tend to be evaluated based on a more traditional organizational role (Johnson et al., 2008). Consequently, males who pursue WLB might be expected to experience greater adverse career progression consequences than females (Frank & Lowe, 2003). On the other hand, more recent studies have suggested that career consequences for males participating in WLB programs like AWAs are less severe than for women (Chung, 2020; Ely & Padavic, 2020).

While WLB program initiatives are predominantly targeted as strategies to retain women employees, they do not necessarily serve as career advancement tools (Almer & Single, 2007). Participation in WLB programs like AWAs by women has been reported as a path toward the effective derailment of their careers (Ely & Padavic, 2020). Therefore, career progression metrics given by supervisors to women who use WLB programs include more of a penalty than that given to similarly situated men. Such differences may help explain why females constitute 47% of large CPA firms, but only 23% of their partners (American Institute of Certified Public Accountants (AICPA), 2019b).

Evidence exists that prevailing views have changed in the marketplace with respect to WLB programs. Responses to a recent survey indicate that there is generally more agreement about the work-life balance advantages of AWAs than about the career consequences to participants in such programs (Knight & Taylor, 2021). A reconciliation of this difference is needed. This study, therefore, can evaluate if the shift to a more supportive view of WLB also manifests in the career progression recommendations by superiors for subordinates.

Once the nuances of career progression are differentiated, the exact nature of their variation by gender are even more challenging to predict. Therefore, the second hypothesis is left without directional prediction and is formally stated as follows:

H2. The effects of work-life balance approaches on career progression will differ between males and females.

2.3. Work-life balance metric (H3)

The transition from a singular all-important objective to a more nuanced multi-dimensional definition of success is one familiar to management in many industries. The first step is usually to elevate the previously neglected aspects of performance to a visible level. Incorporating a non-financial WLB metric in the formal performance evaluation process should be examined as a potential tool for strengthening the effectiveness of audit firm investments in work-life balance and well-being programs.

The implementation of work-life balance and other well-being programs is a significant departure from the traditional work policies and perceptions that previously defined audit firm culture. The effectiveness of such programs is therefore reliant on the reality of organizational change. For change to occur in an organization, traditional approaches must be "unfrozen" so that new methods can be introduced (Levin, 1947). The organizational change model developed by Armenakis and Bedeian (1999) suggests that importance must be placed on conveying an effective change message within the organization. The performance appraisal process within accounting firms constitutes an appropriate medium for communicating desired change from the organization's strategic apex to its lower-level employees.

This study offers a modification to the current performance measurement systems used by accounting firms to incorporate a non-financial measure referenced as WLB. The incorporation of non-financial metrics in performance appraisals has been common practice in the business environment since the introduction of the balanced scorecard (Kaplan & Norton, 1992). The implementation of balanced-scorecard-type performance measurement tools encouraged managers to consider performance of subordinates not only in terms of financial output measures such as net income, but also in terms of actions that indirectly impact an organization's future performance goals. Whereas financial achievements speak for themselves, other measures can more closely reflect a firm's strategic priorities.

The balanced scorecard is a common tool in the managerial process used to develop, communicate, and implement firm strategy (Malina & Selto, 2001). Such a performance measurement tool can be viewed as "the focal point for an organization's efforts" in terms of "defining and communicating priorities" (Kaplan & Norton, 1993). According to organizational change models, using human resource processes such as

the balanced scorecard is an effective method to communicate and incorporate change efforts (Armenakis & Bedeian, 1999).

Although the incorporation of a formal WLB metric in the annual review of employees is not commonplace, there are hints that at least one Big Four accounting firm has evaluated managers on how available they make work/life options for subordinates (Lewison, 2006). Recent discussions with employees of two large public accounting firms (one Big Four and one national firm), however, indicate that while there is an overall firm expectation that WLB should be pursued, it has not yet been equated with more traditional measures of achievement.

The inclusion of a non-financial metric for WLB in the performance management system not only communicates the firm's commitment to change, but also requires that all members of the organization incorporate the concept in their annual review process. This ultimately would lead to goal setting, planning, and other discussions pertaining to this salient strategic change. Furthermore, all members of the organization would be held accountable for the internal firm strategic change as well: employees must set goals as to how they will achieve WLB, supervisors must communicate how they will help subordinates achieve those goals, and ultimately supervisors will be reviewed on how well they support this priority of the organization. Consequently, participant reviewers will be less likely to rely on their traditional thoughts about promotion recommendations that may be inconsistent with the firm's stated goal related to achieving real WLB. In that WLB is believed to increase accountants' work quality (Khavis & Krishnan, 2021), such a change is not farfetched.

Considering firm commitment to WLB as represented by the inclusion of a WLB metric on the performance appraisal tool, supervisors should give higher career progression ratings to individuals who participate in weekend recovery than those who do nothing to achieve WLB. Furthermore, any negative consequences of participation in a WLB program, such as an AWA, is expected to be mitigated when this metric is included on the performance appraisal. However, if supervisors place more value on daily contact than daily availability, an individual who participates in an AWA but communicates with the team at least once on days not in the office may receive higher career progression ratings than an individual who works full-time but does not communicate with the team at all on weekends. While AWA has grown in popularity in the U.S. in recent years (Katz & Krueger, 2019), the willingness of people to work on weekends remains anecdotal and largely invisible.

In general, this study hypothesizes that the WLB metric will improve career progression recommendations for subordinates who take steps to improve work-life balance. In other words, the WLB metric turns AWA and weekend recovery efforts away from the stigma expected in H_1 and toward a career progression enhancement. Those that are doing nothing about WLB would be moved from heroic status to something akin to self-deceiving short run focusing fanatics. However, such a revision of thinking would directly oppose the "ideal worker" prototype. This hypothesis is formally stated as follows:

H3a. When a WLB metric is included on the performance appraisal, individuals who make attempts to achieve WLB will receive stronger career progression recommendations than when a WLB metric is not included.

As previously suggested, the career consequences of attempts to achieve work-life balance may differ by gender. If this is the case in public accounting, the use of a WLB metric on the performance evaluation tool should be expected to mitigate the tendency of reviewers to use gender stereotypes pertaining to the appropriateness of WLB pursuit. Holding supervisors accountable for staff WLB might communicate that it is properly used by both genders. Continuing the logic that gender has always been an explicit dimension of the WLB in accounting literature (e.g., Almer, Cohen, & Single, 2003), the gender implications of a formal metric are hypothesized as follows:

H3b. A WLB metric will mitigate the negative career consequences of WLB efforts equally for males and females.

3. Methodology

This study incorporates a 3 (WLB Approach) x 2 (Gender) x 2 (WLB metric) between-subjects experimental design.[1,2] Participants were randomly assigned to review employee profiles and provide likelihood of recommendation for promotion ratings of the employees and related bonus and salary increase recommendations. In all cases, the hypothetical staff person is described as a well-performing employee. The three manipulated variables or independent variables are WLB Approach, WLB metric, and Gender. The WLB Approach is a categorical variable comprised of three categories. The individual being reviewed in the experiment is described as achieving work-life balance through AWA participation, Weekend Recovery, or making no effort to achieve WLB. The WLB metric relates to whether or not the firm's performance appraisal includes a formal WLB metric. The gender of the subordinate being reviewed is manipulated through the use of unambiguous first names and pronouns to indicate the employee's gender. 50% of participants evaluated a male staff person and 50% of the participants evaluated a female staff person. Qualitative questions were included near the end of the instrument requesting the participant to explain the factors that influenced their recommendations. This data provided additional insight about participant judgments.

3.1. Experimental task

Participants were provided with a summary of performance appraisals for an employee. Participants were first asked to rank how likely they would be to schedule the individual on one of their future jobs. The second question related to how likely the individual would be to receive less challenging work in the future. The first two questions were consistent with the informal feedback asked from participants in the Johnson et al. (2008) study. Participants responded to both questions on a scale of 0 to 10.

After responding to those questions, participants were asked to indicate how likely they are to recommend the employee being reviewed for promotion. This differs from the Johnson et al. (2008) study, in which the formal feedback measure was a rating for performance on a specific project. Rather than rating the subordinate on a single job, participants in our study make a recommendation of whether to promote the subordinate to senior associate, a much stronger career progression decision. For consistency of scale measurement, this also was measured on an 11-point scale ranging from 0 to 10. After ranking likelihood to recommend for promotion, participants had the opportunity to explain their ranking with the use of an open text box. Additionally, participants were asked to provide bonus amount recommendations (between $0 and $10,000) and salary increase recommendations (0% to 10% of base salary).

While an argument could be made to apply factor analysis to the 5 career progression outcomes to produce a more condensed measure for career progression, evaluating each outcome separately provides a

[1] Both authors were associated with the same institution at the time of this experiment. Approval to conduct research using human subjects was obtained from the IRB at the University prior to data collection.

[2] The original experiment incorporated a within subjects design as each participant evaluated 2 subordinates with the same description differing only by gender. There was no significant difference in career progression ratings between the first subordinate reviewed and the second, therefore the results presented in this paper focus on all participants career progression ratings on the first subordinate they reviewed and is therefore essentially data for a between subjects design.

deeper understanding of the multiple dimensions of career progression.[3]

After responding to the task-specific questions, participants were asked to respond to a subset of questions from the Work Arrangement Attitude Scale (Johnson et al., 2012), an approach also consistent with Johnson et al. (2008). The original scale consists of 20 questions that relate to professional/societal ideals, organizational-level outcome issues, and personal-level outcome issues. The scale was developed so that researchers could measure and monitor accounting professional's perceptions of AWAs over time. A subset of this scale was selected to obtain an understanding of manager perceptions related to WLB choices. The measures were combined into a single measure for AWA support by adding the scores of agreement for supportive items and subtracting the scores of the negative or non-supportive items. This measure was then incorporated to all analyses as a covariate (labelled AWAsupport) to control for general participant attitudes toward alternative work arrangements.

Finally, participants were asked to respond to a post-experiment questionnaire which included standard demographic questions as well as questions specific to the participant's personal participation in WLB programs. Additionally, participants were asked to provide information regarding steps/activities their own firms have taken to encourage acceptance and improvement of WLB within their firms.

4. Results

4.1. Sample

This research sought and received the cooperation of the Center for Audit Quality (CAQ) to provide a sample of auditing professionals willing to complete the experimental materials. Although this cooperation underscores the importance of the topic to the public accounting community, it reduces the control of the solicitation by the researchers. Participants included auditors from at least two of the Big Four firms. The researchers do not have precise information about firm affiliation. How many people were approached to create the sample that was attained is also unknown. The data for the experiment was collected between September 2016 and February 2017, prior to the Covid lockdown.

A total of 151 auditors participated in the study. Four submissions were missing a significant number of responses and were deleted. Two additional observations were removed from the analysis due to missing dependent variable measures. Three responses exhibited potential outlier behavior and thus analyses were performed with and without the outliers. No differences in main conclusions were produced by removing these observations. Qualitative responses from those participants were reviewed and since their views are indeed representative of some members of the accounting community, the observations were kept in the dataset. In sum, 145 observations were analyzed to evaluate the hypotheses. Table 1 provides a breakdown of participants by experimental condition as well as demographic information about participants. Means, standard deviations, and correlations between variables are presented in Table 2.

4.2. Analysis

To analyze the hypotheses of the 3 (WLB Approach) x 2 (Gender) x 2 (WLB metric) experiment, a between-subjects GLM Multivariate Analysis of Covariance was performed on five dependent variables associated with career progression of accounting staff: promotion, request, less challenging (inversely coded so that large values represent more challenging assignments), bonus, and raise. IBM SPSS MANOVA Version 28 was used to simultaneously analyze the effects of WLB Approach, Employee Gender, WLB metric, and all two-way interactions on the dependent variables. Adjustment was made for the covariate AWAsupport, a measure intending to capture participants' general attitudes toward alternative work arrangements.

With the use of Wilks' criterion, the combined DVs were significantly affected by WLB Approach, $F(10, 260) = 4.770$, $p < .001$ and the covariate AWA Support, $F(5, 130) = 2.503$, $p < .05$. Neither the main effects of Gender, $F(5, 130) = 0.515$, $p > .05$, nor WLB metric, $F(5, 130) = 0.46$, $p > .05$ significantly contributed to career progression. Similarly, none of the two-way interactions significantly impacted career progression (WLBApp*Gender, $F(10, 260) = 1.49$, $p > .05$; WLBApp*WLBMetric, $F(10, 260) = 1.393$, $p > .05$; Gender*WLBMetric, $F(5, 130) = 0.942$, $p > .05$). Table 3 presents the results of the multivariate analysis.

To investigate the impact of WLB Approach on the Dependent Variables (DVs), univariate and Roy-Bargmann stepdown analysis was performed on the prioritized DVs. The Roy-Bargmann analysis is a procedure that tests the significance of a particular outcome after controlling for all other outcomes preceding it. For this analysis, the potential career progression penalties were ranked based on severity, and then the significance of each DV was analyzed after controlling for all DVs of less severe penalty. Since the dependent variables of career progression are related to each other, this approach accounts for the interdependence among the outcomes. Theoretically, not being assigned challenging work (Less Challenging - inversely coded so that large values represent more challenging assignments) was deemed the lowest level of career progression penalty followed by being less likely to be requested on future work (Request). Next, Bonus, a one-time financial penalty, followed by Raise, a penalty with more permanent consequences, and finally Promotion, the most important marker of career progression.

Since neither of the main effects of Gender and the WLB metric were significant, a more parsimonious model only investigating the effects of WLB Approach on career progression was performed.[4] The remaining independent variables of Employee Gender, WLB metric, and AWA support were controlled as covariates. Univariate F tests indicate WLB Approach significantly influences all career progression measures, with three of the five remaining significant after stepdown. The results indicate there are significant differences among WLB approaches for being assigned less challenging work, $F(2,139) = 6.01$, $p < .01$, Requests on future jobs, $F(2,138) = 8.57$, $p < .001$, and Bonuses $F(2,137) = 8.14$, $p < .001$. Table 4 summarizes the results of this analysis.

A series of ANCOVAs was subsequently performed on each of the three significantly influenced career progression measures (LessChallenging^R, request, bonus) to identify the differences exhibited by the WLB approaches. (See Table 5 Panel A). A summary of the means, standard deviations, adjusted means and standard errors are presented in Table 5 Panel B. Post hoc Tests (Table 5 Panel C) conclude that using an AWA to achieve WLB results in statistically significantly lower career progression measures in comparison to Weekend recovery or making no specific effort to achieve WLB. While average non-finanical career progression scores for those using Weekend Recovery were slightly higher than those making no effort to achieve WLB, the difference was not statistically significant. The financially related career progression scores (bonus, raise) for those using Weekend Recovery were generally lower than those making no effort to achieve WLB, but only bonuses were

[3] Principal components analysis results in a two-component measure of career progression representing financial (Raise and Bonus) and non-financial (LessChallenging^R, Request, Promotion) measures of career progression.

[4] A violation of the homogeneity of variance-covariance matrices was noted. Removal of the potential outliers did not resolve this. Therefore, to ensure robustness of the results, the step-down analysis was re-performed splitting the DVs into two factors (financial and non-financial). The variance-covariance assumption was no longer violated, and results are consistent with original conclusions that WLB Approach significantly impacts career progression in both financial and non-financial measures in the univariate and only non-financial measures after stepdown ($p = .09$).

Table 1
Design and descriptives.

Panel A: Experimental design

		WLB metric		
		Yes	No	Total
WLB approach	None	22	24	46
	AWA	22	26	48
	Weekend	27	24	51
	Total	71	74	145

Panel B: Continuous variables

	Mean	S·D.	Min	Max	N	No response	Total participants
Age	34.59	6.58	27	62	123	22	145
Total experience (in years)	11.45	6.48	4	33	141	4	145
Time with current employer (in years)	9.14	5.24	1	28	141	4	145
Likelihood to participate in WLB in future	4.29	3.13	0	1	105	40	145

Panel C: Discrete variables

Specialization			Gender	N	%
Audit	143	98.62%	Male	80	55.17%
Other	2	1.38%	Female	59	40.69%
Total	145	100.00%	No response	6	4.14%
Big 4			Total	145	100.00%
Yes	79	54.48%	Marital status		
No	66	45.52%	Married	99	68.28%
Total	145	100.00%	Never been married	37	25.52%
Rank			Divorced/Separated	2	1.38%
Manager	62	42.76%	No response	7	4.83%
Senior manager	44	30.34%	Total	145	100.00%
Director	6	4.14%	Children		
Partner	31	21.38%	Yes	72	49.66%
No response	2	1.38%	No	65	44.83%
Total	145	100.00%	No response	8	5.52%
			Total	145	100.00%
Formal WLB metric used by current employer			Participated in a WLB program		
Yes	18	12.41%	Yes	36	24.83%
No	80	55.17%	No	105	72.41%
Do not know	47	32.41%	No response	4	2.76%
	145	100.00%		145	100.00%

Note: Participants could choose not to respond to any question therefore total N for each variable differs.

statistically significantly lower.

Employees in the AWA group were significantly less likely to be given challenging work or be requested on future jobs than individuals who practiced weekend recovery or did nothing to achieve work-life balance. Additionally, employees who chose work over the pursuit of either form of WLB were rewarded with higher bonuses. While this might be more expected relative to those using AWAs, those choosing weekend availability over weekend relaxation were also rewarded with higher bonuses.

On balance, there is considerable support for Hypotheses 1a and 1c. Staff auditors who take AWA will experience slower career progress. Less support exists for Hypothesis 1b. Weekend recovery choices made by staff do not seem to have such distinctive and pervasive career progress consequences. The lack of support for Hypothesis 1b may suggest some shift in cultural acceptance for weekend recovery in a historically traditional "always available" work setting.

Hypotheses 2 focuses on gender effects. Specifically, the results of the GLM analysis in Table 3 indicate that neither the main effect of gender nor the interaction of gender with AWA Approach is significant. Consequently, H2 is not supported. The results suggest there is no difference of career progression scores due to gender of the employee. This perhaps is a historic moment in time for the accounting industry as past literature has consistently suggested a gender gap favoring males in terms of career progression. Furthermore, this is contradictory to findings of Chung (2020) which suggests that males might be less penalized for using AWAs then females in some circumstances. However, gender differences that do exist might be detected with larger sample sizes. Gender differences might also be attributed to those performing the staff evaluation.[5]

The impact of a WLB metric as a tool to mitigate career consequences of taking steps to achieve WLB is the focus of H3. The lack of a significant main effect in the GLM analysis presented in Table 3 suggests that using

[5] The potential for gender differences to exist due to participant gender was also explored. Given the sample size, participant gender effects on the influence of WLB Approach on career progression controlling for AWA support were explored without respect to the employee gender in the experiment. GLM results indicated no significant main effect due to participant gender $F(5,128) = 0.263$, $p > .05$. While the analysis indicated significant interaction with WLB Approach $F(10, 256) = 2.126$, $P < .05$, further investigation with step down analysis followed by ANCOVA with post-hoc analyses found no significant differences in WLB Approach scores due to participant gender. Gender may also be intertwined with firm type if it has produced a disproportionate shift of participants from the large accounting firms to regional or smaller firms for the purpose of gaining more work-life balance. Multivariate analysis was applied to explore differences between participants from large firms and other firms (BIG4). Multivariate analysis indicates potential significant main effects of firm type on career progression while controlling for AWA support, $F(5, 134) = 2.914$, $p < .05$; but not the interaction of WLB Approach and Firm Type, $F(10,268) = 1.221$, $p > .05$.

Table 2
Means, standard deviations, and correlations of dependent and independent variables.

		Mean	S.D	DV 1	DV 2	DV 3	DV 4	DV 5	IV 6	IV 7	IV 8	IV 9
1	Promotion[a]	9.03	1.37	1								
2	Request[b]	8.20	1.47	0.62***	1							
3	LessChallenging[c]	7.69	2.27	0.30**	0.43**	1						
4	Bonus[d]	6.19	2.12	0.30***	0.40***	−0.15	1					
5	Raise[e]	7.03	1.91	0.27**	0.42***	−0.07	0.64***	1				
6	WLB approach	1.03	0.82	0.04	0.02	−0.06	−0.24***	−0.15	1			
7	WLB metric	0.49	0.50	−0.08	−0.04	−0.09	−0.04	−0.07	0.02	1		
8	Gender employee	0.51	0.50	−0.10	−0.07	−0.02	−0.04	−0.03	−0.03	−0.02	1	
9	AWA support	4.06	12.64	0.215*	0.20*	−0.25**	−0.01	0.02	−0.00	−0.15	−0.03	1

a − 11 point Likert scale indicating how likely participant was to recommend the employee for promotion to Senior Associate. (Very Unlikely (0) to Very Likely (10) including Unsure (5).
b- 11 point Likert scale indicating how actively the participant would pursue the opportunity to have the employee work with them on a future client engagement. (Would pursue aggressively (10) to Would not pursue (0) including indifferent (5)).
c-11 point Likert scale indicating the participant's assessment of how likely the employee would be assigned to less challenging work in the future. (Very likely (10) to Very Unlikely (0) including neutral (5). Inversely coded so that higher numbers mean more challenging work.)
d-11 point scale of the amount of discretionary incentive (i.e. bonus) that should be recommended for the employee ($10,000 (10) to $0 (0) each point differing by $1000,
e-11 point scale of the % of base pay salary increase should be recommended for the employee (10% (10) to 0% (0) each point differing by 1%.
WLB = Work-Life Balance.
WLB Approach (None = 0, AWA = 1, Weekend Recovery = 2).
WLB Metric (No metric = 0, Metric provided = 1).
Gender (Female = 0, Male = 1).
*,**,*** - significant at $p < .05$, $p < .01$, $p < .001$, respectively (two-tailed).

Table 3
GLM: Tests of WLB approach, employee gender, WLB metric and interactions multivariate analysis of covariance of career progression.

Effect		Value	df1	df2	F	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	0.983	5	130	1526.82***	0.000	0.983
	Wilks' Lambda	0.983	5	130	1526.82***	0.000	0.983
WLBAPP	Pillai's Trace	0.303	10	262	4.68***	0.000	0.152
	Wilks' Lambda	0.714	10	260	4.77***	0.000	0.155
Gender1	Pillai's Trace	0.019	5	130	0.52	0.765	0.019
	Wilks' Lambda	0.981	5	130	0.52	0.765	0.019
WLBMetric	Pillai's Trace	0.017	5	130	0.46	0.805	0.017
	Wilks' Lambda	0.983	5	130	0.46	0.805	0.017
WLBAPP * Gender1	Pillai's Trace	0.107	10	262	1.49	0.144	0.054
	Wilks' Lambda	0.895	10	260	1.49	0.143	0.054
WLBAPP * WLBMetric	Pillai's Trace	0.100	10	262	1.38	0.189	0.050
	Wilks' Lambda	0.901	10	260	1.39	0.184	0.051
Gender1 * WLBMetric	Pillai's Trace	0.035	5	130	0.94	0.456	0.035
	Wilks' Lambda	0.965	5	130	0.94	0.456	0.035
AWAsup	Pillai's Trace	0.088	5	130	2.50*	0.034	0.088
	Wilks' Lambda	0.912	5	130	2.50*	0.034	0.088

Design: Intercept + WLBAPP + Gender1 + WLBMetric + WLBAPP * Gender1 + WLBAPP * WLBMetric + Gender1 * WLBMetric + AWAsup.
WLB Approach:
None – Hypothetical employee is described as making no effort to achieve a balance between work and life.
AWA – Hypothetical employee is described to participate in an Alternative Work Arrangement with a reduced schedule.
WEREC – Hypothetical employee is described to practice weekend recovery and not responding to emails or calls on weekends.
Gender (0 = Female, 1 = Male).
WLBMetric (No metric = 0, Metric provided = 1).
Yes (1) – The hypothetical employee's summary performance evaluation includes a formal WLB Metric with commentary about WLB efforts.
No (0) – The hypothetical employee's summary performance evaluation does not include any commentary about WLB efforts.
*, ***-significant at $p < .05$ and $p < .001$, respectively (two-tailed).

a WLB metric will have no bearing on mitigating the negative career consequences of WLB approaches. Lack of support for H2 and H3a eliminates the need for further analysis and we must assume that H3b is also unsupported.

The fact that the sample size is not large introduces possible undetected differences for the effect of the WLB metric on each gender. A review of these general relationships based on unadjusted means is presented in Table 6. The basic question is how do audit firms, acting through supervisory personnel, react to differently gendered staff members as they attempt to balance the demands of work and personal lives, and the contributions of a firm sponsored metric for that purpose.

With regards to the people who apparently prioritize always being available for work, the results show superior career progression for females in the absence of a WLB metric. Perhaps in recognition of the larger sacrifice that females are often called upon to make to family life, the results show higher promotion, bonus, and salary recommendations for females. In addition, without a formal WLB metric, females are more likely to be requested for future audit teams and not as likely to be relegated to less challenging work assignments. Adding a formal WLB metric equalizes males and females in terms of being requested on jobs and with bonuses but advances males to a superior position with raises and promotions. With regards to the opposite work-life balance

Table 4
Results of MANCOVA univariate and stepdown analysis of WLB approach on career progression controlling for AWA support, WLB metric, and employee gender.

Panel A: Multivariate tests of significance

	Test	Value	F	Df	Sig
Covariates	Pillai's Trace	0.127	1.21	15, 411	0.262
	Wilks' Lambda	0.875	1.23	14, 373	0.247
WLB approach	Pillai's Trace	0.293	4.66	10, 272	0.000
	Wilks' Lambda	0.723	4.74	10, 270	0.000

Panel B: Univariate and stepdown

IV	DV	*Univariate F*	*df*	*Sig*	*Stepdown F*	*df*	*Sig*
Covariates*	LessChallenging^{R}	3.34*	3, 139	0.021	3.34*	3, 139	0.021
	Request	2.48	3, 139	0.063	1.13	3, 138	0.339
	Bonus	0.36	3, 139	0.781	0.55	3, 137	0.652
	Raise	0.43	3, 139	0.729	0.15	3, 136	0.930
	Promotion	3.09*	3, 139	0.029	1.10	3, 135	0.353
WLB approach	LessChallenging^{R}	6.01**	2, 139	0.003	6.01**	2, 139	0.003
	Request	13.42***	2, 139	0.000	8.57***	2, 138	0.000
	Bonus	13.06***	2, 139	0.000	8.14***	2, 137	0.000
	Raise	4.44*	2, 139	0.014	0.38	2, 136	0.684
	Promotion	6.98**	2, 139	0.001	0.39	2, 135	0.675

R - Inversely Coded so that higher numbers mean more challenging work.
Refer to Table 3 Legend for Variable Definitions.
*,**,*** - significant at $p < .05$, $p < .01$, $p < .001$, respectively (two-tailed).
* *Covariates include Employee Gender, WLB Metric, AWA support.*

approach – requesting a formal alternative work arrangement, the firms are less likely to dispense harsh career consequences to female auditors. Females with AWAs are more likely to be promoted and receive bonuses than are males with AWAs. Along similar lines, such females are more likely to be requested. Females are less likely to be given less challenging work, but when a formal metric is added, the opposite is likely. Similarly, females receive higher raises but when a formal metric is added, males receive higher raises. The advantages enjoyed by women who pursue less formal balance by not working on weekends are less distinct and more mixed. Women and men are mostly equal in their resulting financial remuneration (bonus, raises). Although women have a slight edge in being requested as future team members, they are also more likely to be given less challenging work. Both genders are comparable in being promoted. In general, while not statistically conclusive in this sample, a larger sample may confirm these potential patterns based on staff gender and the potential influence of a formal WLB Metric.

5. Discussion

5.1. Recapitulation and interpretation

Perhaps we should not be surprised at the general tenor of the results. People pay a price when they depart from continuously furthering the impression that they will do anything for their employer. These results suggest that although firms have developed and approved the use of Alternative Work Arrangements as a method to help employees achieve needed work-life balance, choosing to participate in such a program can come at the cost of career progression. AWA participants are less likely to be requested on future jobs, more likely to receive less challenging work in the future, and more likely to be penalized with lower bonuses. Reduced likelihood of challenging work ultimately will probably contribute to even lower likelihood of promotion in the future, since exposure to challenging situations contributes to technical knowledge and the build-out of social networks, factors likely to affect future promotion recommendations. This conclusion is consistent with earlier studies that suggest career progression is hindered by the use of AWAs (Johnson et al., 2012). Supervisors may also be intuitively aware of what meta-analyses have found, that employee effectiveness is inversely proportional to degree of departure from regular work expectations (Baltes, Briggs, Huff, Wright, & Neuman, 1999).

The lack of uniform statistically significant differences between using weekend recovery and making no specific effort to achieve WLB suggests that when being considered for career progression at the staff level, supervisors in public accounting understand the need for recovery and implicitly accept the weekend recovery approach. Although the average career progression measures for weekend recovery participants were lower than individuals who make no effort to obtain WLB, we cannot say that those who refuse to allow their work to become "24/7" will be severely treated when scarce rewards are distributed. Put differently, the marginal reward for those who sacrifice their weekends for work is neither great nor consistent. Even though weekend recovery is not officially recognized by firms as a WLB option, the idea that the weekend is a time when non-work activities *should* be pursued is embedded in the culture. This idea is partially embraced by the supervisory respondents who apparently recognize the value of *not* working during this time.

Unlike most WLB studies in the literature, the present paper did not find pervasive gender effects. That does not mean that gender is not involved. When many dimensions of career progression are simultaneously considered, many individual gender differences emerge. If it is still true that the social obligations placed on women will continue to make AWAs more demanded by women, women may have to realize that there are distinct career progression limits to be expected. However, supervisors in public accounting seem to be somewhat more tolerant of females pursuing WLB strategies. This finding is also consistent with the possibility that men using WLB options may have different connotations (Butler, Gasser, & Smart, 2004).

The failure to find that a balanced-scorecard-type metric legitimizes what most would see as a sane approach to the demanding work of auditors shows us that there is no "magic bullet" to the WLB problem. The demand for hard work exerted over long hours is deeply embedded in the culture of public accounting and apparently cannot be reversed with the formal elevation in importance of WLB according to a new performance evaluation yardstick. The results suggest some resistance

Table 5
Tests of WLB approach on career progression.

Panel A: Summary of ANCOVAS - Effect of WLB approach on career progression

IV	DV	SS	Df	F	Sig	Partial Eta Squared
WLB approach	LessChallenging^{R}	55.662	2, 139	6.01	0.00	0.08
	Request	48.434	2, 139	13.52	0.00	0.16
	Bonus	102.664	2, 139	13.06	0.00	0.16

Panel B: Means, standard deviations, adjusted means and standard errors

		Mean	SD	Adjusted mean	SE
LessChallenging^{R}					
	NONE	7.98	2.54	7.99	0.32
	AWA	6.81	2.54	6.83	0.31
	WEREC	8.25	1.72	8.24	0.30
Request					
	NONE	8.59	1.13	8.59	0.20
	AWA	7.38	1.66	7.39	0.19
	WEREC	8.63	1.23	8.63	0.19
Bonus					
	NONE	7.30	1.90	7.32	0.29
	AWA	5.25	2.07	5.24	0.29
	WEREC	6.04	1.94	6.04	0.28
Raise					
	NONE	7.65	1.95	7.67	0.28
	AWA	6.52	1.80	6.52	0.27
	WEREC	6.94	1.88	6.94	0.27
Promotion					
	NONE	9.26	1.10	9.25	0.19
	AWA	8.46	1.17	8.45	0.19
	WEREC	9.35	1.07	9.35	0.18

Panel C: Group comparisons

	Comparison	Reference	Mean difference
LessChallenging^{R}	AWA	NONE	−1.14*
	AWA	WEREC	−1.44**
	WEREC	NONE	0.30
Request	AWA	NONE	−1.21***
	AWA	WEREC	−1.25***
	WEREC	NONE	0.04
Bonus	AWA	NONE	−2.08**
	AWA	WEREC	−8.16
	WEREC	NONE	−1.27**

R - Inversely Coded so that higher numbers mean more challenging work.
Refer to Table 3 Legend for Variable Definitions.
*,**,*** - significant at $p < .05$, $p < .01$, $p < .001$, respectively (two-tailed).

by supervisors to the WLB criterion, perhaps because of its inconsistency with supervisory self-interest (Powell & Mainiero, 1999). However, this experiment was performed at a single point in time so the full planning, goal setting, educating, and evaluation process that is typical in using a balanced scorecard approach to implement a business strategy throughout an organization could not be replicated in this experiment. A longer-term field experiment that incorporates these aspects of goal planning may result in different conclusions. Therefore, the lack of stronger support for the third hypothesis does not rule out the possibility that a WLB metric can be an effective tool to support WLB efforts.

Qualitative feedback from study participants suggests that the results might have differed regarding promotion and other career progression recommendations if the level of the hypothetical employee had been higher than the staff level that the instrument used. Expectations about constant availability are less tolerant at higher levels within accounting firms. Therefore, WLB efforts, such as taking a sustained break from work on the weekends, might have more negative consequences on career progress at higher ranks within firms.

Perhaps most clearly because it is measured in money, respondents recommending lower bonus amounts may be a reflection of participants' belief that anything less than a "24/7" commitment should equitably result in a proportionately reduced bonus. The singularity of this difference could also reflect a willingness to be less discriminatory across WLB preferences with base level pay than with more discretionary rewards.

The study of WLB should also include the acknowledgement of the "bigger picture." The increase in female labor force participation over the last 50 years is usually interpreted as providing accounting firms with a greatly expanded human capital pool to benefit the servicing of their clientele. However, the conflicting demands placed on this new workforce gradually translated into pressure on employers to create various accommodations that this paper has termed WLB. The failure to respond to this by firms would create a legitimacy crisis, even if doing so would require sometimes expensive recalibrations of expected staff contributions. Institutional theory would predict a formal WLB response that appeared better to external constituents than it would be in

Table 6
General results for gender effects - Unadjusted means.

Career progression measure	Metric	None	Weekend recovery			Alternative work arrangement	
		General results	Mean difference	General results	Mean difference	General results	Mean difference
LessChallenging[R]	No Metric	Female> Male	0.53	Male > Female	0.83	Female > Male	1.31
	With Metric	Female > Male	0.91	Male = Female	0.09	Male> Female	1.17
Request	No Metric	Female > Male	0.60	Female > Male	0.67	Female > Male	0.38
	With Metric	Female = Male	0.09	Female = Male	0.21	Female = Male	0.28
Bonus	No Metric	Female > Male	1.16	Female = Male	0.17	Female > Male	0.62
	With Metric	Female = Male	0.27	Female = Male	0.12	Female > Male	0.55
Raise	No Metric	Female > Male	1.86*	Female = Male	0.00	Female > Male	0.69
	With Metric	Male > Female	1.09	Female > Male	0.68	Male > Female	0.77
Promotion	No Metric	Female > Male	0.44	Feale = Male	0.08	Female > Male	0.62
	With Metric	Male > Female	0.64	Female = Male	0.12	Female > Male	1.47**

R - Inversely Coded so that higher numbers mean more challenging work (i.e. higher career progression).
*, ** Significant at the $p < .05$ and $p < .01$ levels, respectively (two-tailed).

actuality (see Goodstein, 1994). Tapping into the opinions of immediate audit supervisors provides a way to gather insight about what this theory calls the technical core, thereby bypassing what the leaders of these organizations tell us *should* happen regarding WLB. We still seem to be a long way from the flexible-work-for-all accounting world described recently in the *Harvard Business Review* (Donovan, 2019).

Some think that the public accounting work experience will be fundamentally changed for the better by the response to the Covid-19 pandemic (e.g., Gao & Pippen, 2022). Undoubtedly, some professionals have re-evaluated the centrality of work to their lives. However, the lasting effect will more likely change expectations about *how* work is done than about work itself (Thomason & Williams, 2020).

Although this paper has suggested that firms need to be more aware of WLB to maximize the value of their human capital, this does not imply that this perspective is the only purpose of life for workers. The concept of weekend recovery should not be taken as a suggestion that the worker should be exclusively devoted to the work that is to be done thereafter. Recovery is a byproduct of the many joys available with the discretionary time to pursue them.

5.2. Contribution

This study adds to the WLB literature in public accounting by exploring the impact of WLB approach choices on supervisor perceptions and subsequent decisions about career progression. Furthermore, this study extends and builds on the current body of literature that explores the impact of alternative work arrangements (AWAs) on careers in public accounting. Specifically extending the Johnson et al. (2008) study that focused on the perceived career consequences of AWA participation, the present research suggests AWA participation, more than a decade later, is still problematic. Perhaps this result is not surprising in an era wherein professional firms are increasing financial rewards to those willing to work extraordinary hours (Cain Miller, 2019). Furthermore, this study considers the use of weekend recovery efforts as a readily available approach to WLB not previously evaluated in the accounting literature.

Career progression implications were explored in this study by asking experienced participants to provide separate recommendations for promotion, bonuses, and salary increases of a subordinate. Rather than just allowing participants to self-determine the value of WLB, a formal appraisal tool was proposed as a method to communicate the firm's official position about the importance of WLB for staff. Both features represent enhancements for the study of performance evaluation in public accounting.

One of the major advantages of this study is its robust view on career facilitation. Often recommendations for salary, promotion, bonus or the chance to prove what one can do with meaningful work, operate in different ways and send different signals. Research on just one dimension cannot be taken as a proxy for other dimensions. Career progress is a highly nuanced affair with many vectors, only some of which are financial and immediate. Being wanted for teams and being given important growth opportunities provide a staff person with invaluable long-term professional enhancement. Only by including several of these dimensions can subtle performance evaluation differences, even among those that are primarily financial, be detected.

This paper's results suggest that work-life balance concerns do not necessarily work to the disadvantage of female staff in accounting. Women that seek WLB in their lives are not disproportionately penalized for that choice. That said, the individual decisions that are made in performance evaluation are far from gender neutral. Large sample studies are needed to ascertain the true magnitude of these differences.

According to management theory, building strategic change into human resource processes should encourage internal change in beliefs within firms. Here, the implicit expectation was that a modification to formal appraisal tools would ultimately deepen supervisor perceptions of the value of WLB program use. This should make WLB programs that are already provided by firms more effective for retention purposes, and perhaps motivate supervisors to rethink the level of demands being made on staff for their time. This study that explores the impact of adding an explicit WLB metric to formal performance appraisals demonstrates that culture change is not easy. Without full support by immediate supervisors, firm level WLB programs will not be well aligned with the more pressing reality of performance evaluation. Such an alignment will require a better understanding of how people can attain such balance in their lives. This might force a reconsideration of how work tends to spill over, often via modern technology, into time traditionally reserved for personal and family pursuits.

5.3. Limitations and opportunities for future research

This study shares all the shortcomings of experimental work in the accounting discipline. Presenting hypothetical information to people always involves a leap of faith that they will respond in a way consistent with their response to real people and real situations. Although manipulation checks contained in the instrument produced ample evidence that respondents understood the information they were given, understanding and embracing the importance of the scenarios are not always the same. However, the use of real auditors with requisite levels

of performance evaluation experience minimizes the likelihood that respondents were, as put by Nisbett and Wilson (1977), telling us more than they could really know.

The manipulations of this experiment were limited to the WLB approach chosen and the inclusion of a formal WLB metric. Future research might consider the manipulation of family structure and/or commitments of the employee being evaluated. While reviewer family structure did not seem to affect the decisions that reviewers in this experiment made about their subordinates, knowledge of the reviewee's family commitments may have an impact. Future research might also explicitly consider gender identity.

Behavioral work could always benefit from a larger sample size. Here, one could think that the sample size might be too small to detect a benefit from the WLB metric used. The inability to simulate the planning and goal setting process typically involved in using balanced scorecards may have been a larger hinderance than originally presumed in finding the tool effective. A field experiment over time might produce different results. Small sample size also limits the number of covariates that can be included in such analysis limiting the ability to further explain the lack of effect of the WLB metric. The non-significance of this instrument in performance evaluations does not further the general idea that making things visible changes attitudes.

Although the recent Covid pandemic has been widely believed responsible for an acceleration of the idea that more flexibility is needed about how work is conducted, it remains doubtful that this will reverse the negative career consequences of AWAs and for unlimited WLB demands. More likely is that perceptions about the role of work may eventually adjust to a level permitting a WLB metric to become more effective. The forced acceptance of AWAs for both genders during COVID may result in fewer future gender implications. Without diminishing the rapid and impressive progress toward neutrality that has been made by public accounting, we would still bet on the immutability of gendered expectations. The influx of females has not changed the fact that the "ideal worker" in public accounting has always been male and that any special allowances to others should be made less than willingly (Hardies, Lennox, & Li, 2018). As Lightbody (2008) concludes, flexible work and WLB does not change accounting culture, it just enables it.

Funding for the research project described in this article was provided by the Center for Audit Quality. However, the views expressed in this article and its content are those of the authors alone and not those of the Center for Audit Quality.

Declaration of Competing Interest

Neither of the authors has a conflict of interest to report, Tim Fogarty.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to extend thanks to Dennis Caplan, the Center for Audit Quality, the participants in this study, Jagdip Singh, Bob Libby, Chris Smith, Celeste Suchko, the Department of Accountancy at Case Western Reserve University, and the reviewers of the 2018 AAA Annual Meeting and Conference and Ohio Region Meeting.

References

Adamson, G. (2013). Here comes the baby Boomer bubble – Valuing your practice for partner retirements. *CPA Practice Advisor*, (Feb 14).

Almer, E. D., & Single, L. E. (2007). Shedding light on the AICPA work/life and Women's initiatives research: What does it mean to educators and students? *Issues in Accounting Education, 22*(1), 67–77.

Almer, E. D., Cohen, J., & Single, L. (2003). Factors affecting the choice to participate in flexible work arrangements. *Auditing: A Journal of Practice & Theory, 22*(1), 69–91.

Almer, E. D., & Kaplan, S. E. (2002). The effects of flexible work arrangements on stressors, burnout, and behavioral job outcomes in public accounting. *Behavioral Research in Accounting, 14*(March), 1.

American Institute of Certified Public Accountants (AICPA). (2004). *AICPA work/life initiatives 2004 research – A decade of changes in the accounting profession: Workforce and Women's trends and human capital practices*. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). (2011). *The 2011 PCPS top talent study*. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). (2015). *2015 trends in the supply of accounting graduates and the demand for public accounting recruits*. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). (2019a). *PCPS CPA firm top issues survey commentary*. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). (2019b). *Trends in the supply of accounting graduates and the demand for public accounting recruits*. New York, NY: AICPA.

Anderson, J. C., Johnson, E. N., & Reckers, P. M. J. (1994). Perceived effects of gender, family structure, and physical appearance on career progression in public accounting: A research note. *Accounting, Organizations and Society, 19*(6), 483–491.

Anderson-Gough, F., Grey, C., & Robson, K. (2000). In the name of the client: The service ethic in two professional services firms. *Human Relations, 53*(9), 1151–1174.

Armenakis, A. A., & Bedeian, A. G. (1999). Organizational change: A review of theory and research in the 1990s. *Journal of Management, 25*, 293–315.

Baltes, B. B., Briggs, T., Huff, J., Wright, J., & Neuman, G. (1999). Flexible and compressed workweek schedules: A meta-analysis of their effects on work-related criteria. *Journal of Applied Psychology, 84*(4), 496–513.

Barry, J. S. (2015). A sustainable profession. *CPA Journal, 85*(4), 5–6.

Binnewies, C., Sonnentag, S., & Mojza, E. J. (2010). Recovery during the weekend and fluctuations in weekly job performance: A week-level study examining intra-individual relationships. *Journal of Occupational and Organizational Psychology, 83*(2), 419–441.

Boomer, J. (2018). The state of accounting firm succession. *CPA Practice Advisor*, (July 10).

Buchheit, S., Dalton, D. W., Harp, N. L., & Hollingsworth, C. W. (2016). A contemporary analysis of accounting Professionals' work-life balance. *Accounting Horizons, 30*(1), 41–62.

Butler, A., Gasser, M., & Smart, L. (2004). A social-cognitive perspective on using family-friendly benefits. *Journal of Vocational Behavior, 65*(1), 57–70.

Cain Miller, C. (2019). Women did everything right. Then, work got greedy. *The New York Times*, B1–B7. April 29.

Chung, H. (2020). Gender, flexibility stigma and the perceived negative consequences of flexible working in the U.K. *Social Indicators Research, 151*, 521–545.

Cohen, J., Krishnamoorthy, G., & Wright, A. (2010). Corporate governance in the post-Sarbanes-Oxley era: Auditors' experiences. *Contemporary Accounting Research, 27*(3), 751–786.

Cohen, J. R., & Single, L. E. (2001). An examination of the perceived impact of flexible work arrangements on professional opportunities in public accounting. *Journal of Business Ethics, 32*(4), 317–328.

Cooper, C. L., Davidson, M. J., & Robinson, P. (1982). Stress in the police service. *Journal of Occupational Medicine.: Official Publication of the Industrial Medical Association, 24*(1), 30–36.

Davidson, J. (2018). 6 steps to work-life balance for CPAs. *Text. AccountingWEB. August, 31*, 2018. https://www.accountingweb.com/practice/growth/6-steps-to-work-life-balance-for-cpas.

Donovan, A. (2019). What PwC learned from its policy of flexible work for everyone. *Harvard Business Review*, 71–74. January 29.

Ely, R. J., & Padavic, I. (2020). *What's really holding women back?* Harvard Business Review. March/April.

Fogarty, T., Parker, L., & Robinson, T. (1998). Where the rubber meets the road: Performance evaluation and gender in large public accounting organizations. *Women in Management Review, 13*(8), 299–310.

Fortune. (2021). 100 best companies to work for. https://fortune.com/best-companies/2021/.

Frank, K. E., & Lowe, D. J. (2003). An examination of alternative work arrangements in private accounting practice. *Accounting Horizons, 17*(2), 139–151.

Gao, S., & Pippen, S. (2022). The accounting profession and the pandemic: Temporary and long term changes. *CPA Journal*, 32–36. February.

Goodstein, J. D. (1994). Institutional pressures and strategic responsiveness: Employer involvement in work-family issues. *Academy of Management Journal, 37*(2), 350–382.

Gullapalli, D. (2005). Take this job and… file it; burdened by extra work created by the Sarbanes-Oxley act, CPAs leave the big four for better life. *Wall Street Journal*, C1. May 4.

Hardies, K., Lennox, C. S., & Li, B. (2018). Gender discrimination? Evidence from the Belgian public accounting profession. *Evidence from the Belgian Public Accounting Profession*. https://doi.org/10.2139/ssrn.3258719. Available at SSRN:.

Hart, M. (2021). Protect staff's mental health with true time off. *Journal of Accountancy, 232*, 73–77. December 17.

Jain, S., & Nair, S. K. (2013). Research on work- family balance: A review. *Business Perspectives and Research, 2*(1), 43–58.

Johnson, E. N., Jordan Lowe, D., & Reckers, P. M. J. (2008). Alternative work arrangements and perceived career success: Current evidence from the big four firms in the US. *Accounting, Organizations and Society, 33*(1), 48–72.

Johnson, E. N., Lowe, D. J., & Reckers, P. M. J. (2012). Measuring accounting Professionals' attitudes regarding alternative work arrangements. *Behavioral Research in Accounting, 24*(1), 47–71.
Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – Measures that drive performance. *Harvard Business Review*, 71–79 (January–February).
Kaplan, R. S., & Norton, D. P. (1993). Putting the balanced scorecard to work. Harvard Business Review. *(September-October)*, 24–35.
Katz, L. F., & Krueger, A. B. (2019). *Understanding trends in alternative work arrangements in the United States* (Vol. No. w25425). National Bureau of Economic Research.
Kezar, A. J. (2001). Understanding and facilitating organization change in the 21st century. In , *Vol. 28, Number 4. ASHE-ERIC higher education report*. San Francisco: Jossey-Bass.
Khavis, J., & Krishnan, J. (2021). Employee satisfaction in accounting firms, work-life balance, turnover, and audit quality. *Auditing: A Journal of Practice & Theory, 40*, 161–192.
Knight, M. E., & Taylor, R. M. (2021). A word of caution to students: Public accountant perceptions of alternative work arrangements. *Issues in Accounting Education, 36*(3), 29–47.
Kornberger, M., Carter, C., & Ross-Smith, A. (2010). Changing gender domination in a big four accounting firm: Flexibility, performance, and client service in practice. *Accounting, Organizations and Society, 35*(8), 775–791.
Ladva, P., & Andrew, J. (2014). Weaving a web of control: 'The promise of opportunity' and work-life balance in multinational accounting firms. *Accounting, Auditing & Accountability Journal, 27*(4), 634–654.
Levin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations, 1*, 5–41.
Lewison, J. (2006). The work/life balance sheet so far. *Journal of Accountancy, 202*, 45–49. August.
Lightbody, M. (2008). "No time to stand and stare" imagery of flexible work arrangements in Australian professional accounting journals. *Pacific Accounting Review, 20*(2), 138–157.
Lightbody, M. G. (2009). Turnover decisions of women accountants: Using personal histories to understand the relative influence of domestic obligations. *Accounting History, 14*(1–2), 55–78.
Little, L. (2018). How to find a work-life balance as an accountant. AccountancyAge.com *https://www.accountancyage.com/2018/05/02/how-to-find-a-work-life-balance-as-an-accountant/*.
Malina, M. A., & Selto, F. H. (2001). Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research, 13*(September), 47–90.
Mendlowitz, E. (2018). State of the profession 2018: Views on practice management. *CPA Journal, 88*(12). December.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231.
Ono, Y., Watanabe, S., Kaneko, S., Matsumoto, K., & Miyao, M. (1991). Working hours and fatigue of Japanese flight attendants (FA). *Journal of Human Ergology, 20*(2), 155–164.
PCAOB. (2013). Standing advisory group meeting discussion. In *Audit quality indicators*. May 15–16.
Pelletier, L. G., & Vallerand, R. J. (1996). Supervisors' beliefs and Subordinates' intrinsic motivation: A behavioral confirmation analysis. *Journal of Personality and Social Psychology, 71*(2), 331–340.
Pencavel, J. (2015). The productivity of working hours. *The Economic Journal, 125*(589), 2052–2076.
Powell, G. N., & Mainiero, L. A. (1999). Managerial decision making regarding alternative work arrangements. *Journal of Occupational and Organizational Psychology, 72*(1), 41–56.
PricewaterhouseCoopers (PwC). (2013). *PWC's next gen: A global generational study*. Evolving Talent Strategy to Match the New Workforce Reality.
PwC. (2020). Building on a culture of belonging. https://www.pwc.com/us/en/about-us/newsroom/press-releases/building-on-a-culture-of-belonging-diversity-inclusion.html.
Ragsdale, J. M., & Beehr, T. A. (2016). A rigorous test of a model of employees' resource recovery mechanisms during a weekend. *Journal of Organizational Behavior, 37*(6), 911–932.
Rudman, L. A., & Mescher, K. (2013). Penalizing men who request a family leave: Is flexibility stigma a femininity stigma? *Journal of Social Issues, 69*(2), 322–340.
Schulte, B. (2015). *Overwhelmed: How to work, love and play when no one has the time*. New York: Picador.
Sweeney, J. T., & Summers, S. L. (2002). The effect of the busy season workload on public accountants' job burnout. *Behavioral Research in Accounting, 14*(1), 223–245.
Thomason, B., & Williams, H. (2020). What will work-life balance look like after the pandemic. *Harvard Business Review*, 1–4. April 16.
Thornton, G. (2015). Grant Thornton introduces industry-leading flex time off policy for employees. *Press Release*, (September 30).
Twenge, J. M. (2010). A review of the empirical evidence on generational differences in work attitudes. *Journal of Business and Psychology, 25*(2), 201–210.
Tysiac, K. (2017). Flexibility as a retention tool. *Journal of Accountancy, 232*(December), 32–34.
Virtanen, M., Heikkilä, K., Jokela, M., Ferrie, J. E., David Batty, G., Vahtera, J., & Kivimäki, M. (2012). Long working hours and coronary heart disease: A systematic review and Meta-analysis. *American Journal of Epidemiology, 176*, 586–596.
Williams, J. (2000). *Unbending gender: Why family and work conflict and what to do about it*. New York: Oxford University Press.
Williams, J. C. (2010). *Reshaping the work-family debate: Why men and class matter*. Cambridge, MA: Harvard University Press.
Williams, J. C., Blair-Loy, M., & Berdahl, J. L. (2013). Cultural schemas, social class, and the flexibility stigma. *Journal of Social Issues, 69*(2), 209–234.