



文章编号: 1005-9679(2017)01-0068-05

基于推特数据的“工作-家庭冲突”时序模式研究

刘义理

(同济大学管理科学与工程系, 上海 200092)

摘要: 工作-家庭冲突是组织行为学和人力资源管理领域的重要研究课题, 以往成果中冲突与工作满意、家庭满意和生活满意的联系本质与强度存在较大差异, 需要更为合理的事实观测加以解释。本文采用大数据分析, 对美国推特用户的工作-家庭冲突模型进行了时序观测和分析。结果显示, 较长时间内美国人的冲突程度趋于平稳, 波动主要存在于以星期为单位的周期内, 基于时间的冲突对工作满意度有直接影响, 其他情况下冲突与满意度之间没有显著联系。使用截面数据的传统问卷调查法是造成差异的可能原因。

关键词: 工作-家庭冲突; 大数据分析; 推特
中图分类号: F063.4 **文献标志码:** A

1 研究背景

工作-家庭冲突 (Work-Family Conflict) 是组织行为学 (OB) 和人力资源管理 (HRM) 领域日趋重要的研究论题, 该主题研究历时 30 年, 始终维持稳定增长。工作-家庭冲突问题由于会引发员工不满、离职率攀升、工作压力增加等诸多消极结果, 并对组织绩效、心理契约、组织承诺等产生影响, 甚至影响组织战略制定。

迄今为止, 工作-家庭冲突方面的研究存在着几个主要问题。首先, 一直以来, 研究者都认为工作-家庭冲突使得员工的工作和家庭满意度较低, 影响身体和精神健康, 导致较高的离职率和工作压力。但是这些联系的本质和强度在不同的研究中却大不相同, 从微不足道到强的各种结论均有呈现, 学者们认为有必要对于这种差异做出解释, 然而目前这方面的研究却非常缺乏。此外, 很少有研究关注工作-家庭冲突本身不断变化的本质。最近的研究认为对工作-家庭冲突的感知依赖于个人的特点和经验, 不同的人对于冲突和冲突产生的影响会有着不同的感受。这意味着人们可以不断适应冲突, 或者能够更加有效地处理这种冲突。因此, 应该采用更加动态而非静态的观点来研究工作-家庭冲突问题。由此, 一系列相应的基础性问题亟需做出回答: 工作-家庭冲突的进化趋势是什么? 工作-家庭冲突与工作满

意度和家庭满意度的关系是否依然存在? 相应的, 工作和家庭满意是否还是生活满意度的重要组成部分?

理论上, 对工作-家庭冲突的时序模式的观察将有助于解决这些问题。然而在本研究领域中占统治地位的调查问卷法, 大都使用截面数据而非面板数据, 不能有效支持时序模式的分析。除此之外, 问卷方法还存在诸多局限: 受访者可能不理解问卷的问题; 有限的选项不能准确反映受访者的意见; 受访者通常需要通过回忆来答题, 造成作答不准确或不全面, 即回忆偏见; 大多数调查使用同质性的人群, 规模较小 (通常几百人, 很少的时候千人以上), 样本的选择限制了研究成果的普遍性价值。为此, 亟需选用更为合适的研究方法。

作为当下最流行的社交媒体网站之一, 推特 (Twitter) 为这方面研究提供了更好的机会, 有助于克服上述研究方法的局限性。推特用户自主地记录和表达着他们的生活内容, 体现了他们的真实感受和情感; 推特内容的生成并不由少数的问题触发, 不存在误解或者不明白的情况, 而且对生活经历的记录更加全面和完整, 不受少数选项的限制; 通常推特发帖是对当时或短时间内用户生活和情感的快照, 可以大大减少回忆偏见的产生。各种各样用户发布的数以亿万计的帖子, 同

基金项目: 国家自然科学基金, 71101110, 71202034。

作者简介: 刘义理, 博士, 同济大学经济与管理学院讲师。



时具有更细的粒度和更大的规模，将能够有效提高研究成果的普适性价值。

为此，本研究使用美国用户的推特数据进行工作 - 家庭冲突在时间维度上的模式变化研究，检查大数据背景下工作 - 家庭冲突和工作满意、家庭满意及生活满意的关联情况。本研究既可以对未来工作 - 家庭冲突的研究提供基础性支持，也有助于数据分析方法在具体管理领域的扩展和深入研究。包括本部分的研究背景介绍，本文在第二部分进行了相关研究的综述，第三部分介绍了数据收集及数据处理的方法及过程，第四部分为结果分析及解释，最后做出了研究结论与未来研究的展望。

2 数据处理

2.1 数据预处理

本研究使用的推特数据集包含美国 121022 个用户在 2009 年 9 月到 2010 年 1 月所发的 9000659 条帖子，推特的用户样本很好地代表了美国实际的人口分布。

因为工作 - 家庭冲突研究需要针对有工作的推特用户，根据用户发帖主题进行了筛选，去除帖子内容从不涉及工作的用户；为了更好地研究，也过滤掉发帖总数不超过 10 条的用户。处理后的样本数据为：用户数 -93,700，总帖数 -3,327,715，人均发帖 -35。此外，将发帖时间从标准时间(UTC)转换为发帖者的本地时间。

我们使用 LIWC(2007, 语言学查询及文字统计软件)进行帖子的处理。该软件使用典型的文字字典来区分 7 类个人相关的活动主题(比如，工作、家庭、住所、休闲等)，32 个文字类型来区分心理活动(比如，情感、认知和生理过程)和其他语言学类别。

2.2 指标设计

本研究设计了四个变量用以研究的测度：工作 - 家庭冲突，工作满意度，家庭满意度和生活满意度。

基于以前的研究，我们使用两个方法来测量工作 - 家庭冲突：基于时间的冲突和基于压力的冲突。

基于时间的冲突 (Time-based Conflict)：是人们在不同领域中对时间的竞争结果。典型的问题是“我在家庭(工作)上花费的时间导致我没有更多的时间用于工作(家庭)”。我们使用一天中不同主题的帖子在所有发帖中所占的比例来测量人们在这方面的时间花费。具体而言，基于时间的冲突可以定义为：

$$TC_i(d) = \frac{TA_i(d) - BTA_i}{BTA_i} \quad (1)$$

其中

$$TA_i(d) = \frac{Tweets_{iW}(d)}{Tweets_i(d)} / \frac{Tweets_{iF}(d)}{Tweets_i(d)} = \frac{Tweets_{iW}(d)}{Tweets_{iF}(d)} \quad (2)$$

$$BTA_i = \frac{1}{\|D_i\|} \sum_{d \in D_i} TA_i(d) \quad (3)$$

式中 $i \in I, d \in D_i$,

I 表示所有用户的集合， D_i 为人们发帖的日期集合。 $Tweets_{iW}$ 表示用户 i 所发的有关工作(Work)的帖子数， $Tweets_{iF}$ 表示用户 i 所发的有关家庭(Family)的帖数。 TA_i 表示工作和家庭领域的时间分配比率， BTA_i 为个人在不同领域的时间分配基线，它移除了个人在不同时间里的过高或过低的波动，定义中“更多”即表现为超出日常的时间分配。这样， $TC_i(d) > 0$ 就表示用于工作的时间超出正常状态下人们在这两个领域上的时间分配比例，由工作到家庭的基于时间的冲突就会发生。反之，如果 $TC_i(d) < 0$ ，就会产生从家庭到工作的基于时间的冲突。

群体级别上的基于时间的冲突可以定义为：

$$ATC(d) = \frac{1}{\|I(d)\|} \sum_{i \in I(d)} TC_i(d) \quad (4)$$

其中 $I(d)$ 为用户集合的子集，表示 d 天的活跃用户集。

基于压力的冲突 (Stress-based Conflict)：是指个人在一个角色上的压力引起另一个角色的紧张，阻碍另一个角色上的期望实现。压力可以从心理学和身体反应的角度加以解释。心理学上的压力表现包括紧张、焦急、脆弱、压抑、冷漠和易怒，身体反应的表现主要是人们的行为和活动。因此，需要从这两个方面对基于压力的冲突进行测量。

情绪能够同时反映人们的心理和身体压力。具体而言，由于工作 - 家庭冲突而产生消极情绪体现为对于工作 / 家庭的不满。消极的情绪将导致人们在一个领域中悲观，行为和活动的效率低下。所以，可以使用消极情绪作为测量压力的指标。基于压力的冲突 (SC) 可以定义为：

$$SC_{iW \rightarrow F}(d) = \begin{cases} \frac{NS_{iF}(d)}{NS_{iW}(d)}, & \text{if } (NS_{iW}(d)) \neq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

$$SC_{iF \rightarrow W}(d) = \begin{cases} \frac{NS_{iW}(d)}{NS_{iF}(d)}, & \text{if } (NS_{iF}(d)) \neq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

其中

$$NS_i(d) = \frac{S_i(d) - BS_i}{BS_i} \quad (7)$$



$$S_i(d) = \frac{Tweets_negative_i(d)}{Tweets_i(d)} \quad (8)$$

$$BS_i = \frac{1}{||D_i||} \sum_{d \in D_i} S_i(d) \quad (9)$$

$S_i(d)$ 表示用户 i 在天 d 中工作和家庭领域内的消极情绪。 BS_i 表示个人用户的消极情绪基准, $NS_i(d)$ 表示经过标准化处理后的消极情绪值。群体指标为:

$$ASC(d) = \frac{1}{||I(d)||} \sum_{i \in I(d)} SC_i(d) \quad (10)$$

这里使用积极情绪来测量人们的工作、家庭和生活满意度。满意可以解释为情绪对一个领域内绩效和回报的影响结果。具体来说, 积极的情绪可以提高认知功能, 提高工作积极性和坚持度, 提升与他人的积极互动。所有这些产出可以改善人们的角色性能。同样的, 由于角色性能的改善带来的内在和外在的回报可以提高人们的情绪。特别说明的是, 与直觉不同, 积极情绪和消极情绪并不总量守恒, 即积极情绪升高必然导致消极情绪下降, 反之亦然, 两者是相对独立的变量, 因而可以分别用作不同的测量指标。

这里使用相关主题 (工作, W ; 家庭, F ; 和生活, L) 的积极情绪帖来测度三种满意度:

$$S_i(k, d) = \frac{Tweets_{i_p}(k, d)}{Tweets_i(k, d)} \quad (11)$$

其中 $k \in K, K = \{W, F, L\}$.

同上, 我们使用标准化的指标来描述三种类型的群体满意度

$$AS(k, d) = \frac{1}{||I(d)||} \sum_{i \in I(d)} \frac{S_i(k, d) - BS_i(k)}{BS_i(k)} \quad (12)$$

其中

$$BS_i(k) = \frac{1}{||D_i||} \sum_{d \in D_i} S_i(k, d) \quad (13)$$

2.3 测量方法

对于研究提出的第一个问题: 工作-家庭冲突的趋势是什么, 更加严重还是减缓? 需要观察持续时间里两种方式冲突的时序变化模式, 即 $ATC(d)$, $ASC_{W \rightarrow F}(d)$ 和 $ASC_{F \rightarrow W}(d)$ 的时序变化。

对于第二个问题: 工作-家庭冲突对工作、家庭和生活满意度是否还具有明显的影响? 需要观察相应变量对的时序变化模式: $AS(k, d)$ 和 $ATC(d)$, $AS(k, d)$ 和 $ASC_{W \rightarrow F}(d)$ 。

对于第三个问题: 工作满意度和家庭满意度是否还是生活满意度的重要原因, 需要观察变量对变化: $AS(F, d)$ 和 $AS(L, d)$, $AS(W, d)$ and $AS(L, d)$ 之间的关联关系。

3 结果与分析

3.1 工作-家庭冲突的趋势变化

简单地观察冲突指标在 5 个月内的变化并不能反映实际的情况, 因为人们会在周中 (工作日)

和周末 (休息日) 呈现不同的状态。因而这里将研究的周期单元定为一个星期, 具体观测多个星期中的变化规律。数据分析显示, 不同星期的时序变化模式和观测值基本相同, 并没有太大的差异。表示在这个时间之内, 美国人的工作-家庭冲突状态基本平稳, 没有明显加剧或者趋缓。

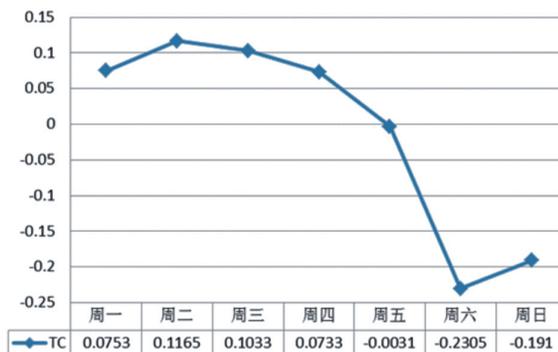


图 1 TC 的时序变化

图 1 显示出基于时间的冲突的时序变化模式。可以看到, 从周一到周四, $TC > 0$, 表示人们在这些日子里花在工作上的时间要超过花在家庭领域的时间。时间花费在周二达到高潮, 而后逐渐下降。从周五到周日, $TC < 0$, 也就是说, 人们开始将时间更多地花费在家庭领域, 有趣的是, 这种转变从周五就开始了, 而不是在周末才有。从绝对值的大小来看, 周末人们在两个领域的时间差要远大于平时周中的时间差。

因此, 如果采用问卷调查的方法, 在考察基于时间的冲突时, 受访者在不同时间给出的答案可能会受到当时情况的影响, 容易有较大的变化, 从而导致研究产生不同的观测结果。

图 2 显示了基于压力的冲突每天的变化规律。数据显示, 工作上的压力与家庭生活的压力存在正相关关系。

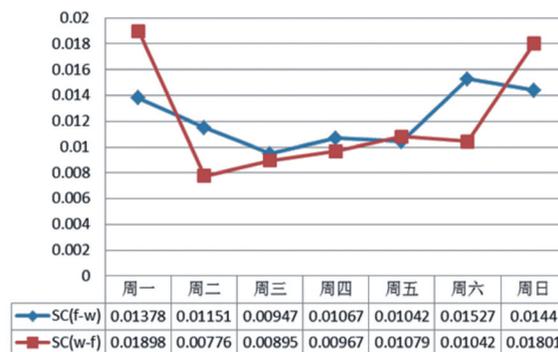


图 2 SC 的时序变化

与基于时间的冲突不同, 工作-家庭方向的压力冲突在周一和周日达到高点, 这可能是因为周日和周一处于工作和家庭休息的临界点, 状态的转换导致人们的适应程度产生了较大的波动。工作领域中微小的消极情绪会更多地影响家庭生活的消极



情绪。这种情形也发生在家庭 - 工作方向的基于压力的冲突模式中。可以看到, 虽然程度略有降低, 周一和周日的家庭 - 工作方向的冲突值也处于高点, 表示家庭领域中的消极情绪会更多地影响人们在工作中的情绪。与前者不同的是, 家庭 - 工作方向的冲突在周六达到高点, 这意味着整个周末, 家庭生活成为美国人的重要影响因素, 家庭生活上的不愉快将极大地影响周末对工作的卷入。

在周中的日子里, 两个方向的冲突呈现相对平稳的状态, 更低的值表示这段时间内, 一个领域内的压力对另一个领域的压力影响相对较小。

在周三到周五, 两个方向的冲突值基本相同, 意味着人们对工作领域和家庭领域的重视程度基本相同, 周二和周六, 家庭领域的消极情绪更容易导致工作领域的消极情绪, 而在周一和周日, 情况正好相反, 工作领域的消极情绪更大地影响着家庭领域的消极情绪。

有趣的是, 在周一和周日, 工作上的消极情绪对家庭领域有着更多的影响, 相比而言, 周二和周六家庭领域的消极情绪对工作领域有着较多的影响。

无论是单独考察某个方向的冲突, 还是两个方向冲突的相加, 不同日子里人们对这些指标的反应并不相同, 存在较大的差异。与基于时间的冲突的时序模式类似, 问卷采集的答案很可能依然会在不同的时间点反映当时人们的认知, 从而产生不同的回答。这可以从一方面解释为什么不同的研究中得到的工作 - 家庭冲突的情形和强度各不相同。然而, 将这些不同放在更久的时间范围内进行观察, 就会发现, 实际上人们在很长一段时间内对于生活中的工作 - 家庭冲突状况并没有太大的波动, 这可能是因为人们已经习惯了由于 IT 技术、互联网和通信技术带来的工作领域和家庭领域的界限模糊及重叠, 已经能够接受当下的生活模式并泰然处之。对于国家和企业的管理者来说, 采取过度敏感的对策其实并不必要。

3.2 冲突及满意度的联系

首先来看基于时间的冲突与三种满意度的周期变化图。

从图 3 可以看出, 从周一到周四, 随着 TC 的变化, LS 和 FS 并没有明显的效果, 但是 WS 与 TC 有类似的形式。具体来说, 生活满意度 LS 呈现相对平稳的状态, 在周末更是有稍微的上升, 这表明时间在工作和家庭两个领域间的分配变化并不显著影响生活满意度; 相类似的, 家庭满意度 FS 也呈现相对平稳的状态, 反映出时间分配的多少并不显著影响人们对家庭生活的满意。这说明, 随着 IT 的渗透逐步普遍和深入, 人们已经不再单纯以时间的多少来衡量家庭满意和生活满意, 生活的内容等其他因素可能对美国人更为重要。

但是工作满意度 WS 则与 TC 呈现相对同步的规律, 周中在工作上分配的时间较多, 工作满意度也较高, 而到了周末, 在家庭上花费的时间更多, 工作满意度降低。此外, 使用回归分析检验 TC 与 WS、FS 和 LS 的关系, 结果显示, $WS(<0.001)$ 确实与 TC 有很强的关联关系。考虑美国的文化特点, 可以解释为美国人并不喜欢在私人的休闲时间(比如周末, 通常作为家庭生活的的时间)继续工作, 即使只是占用比较少的时间也是如此。但是这个规律在其他文化背景的国家是否成立还有待更多观测。

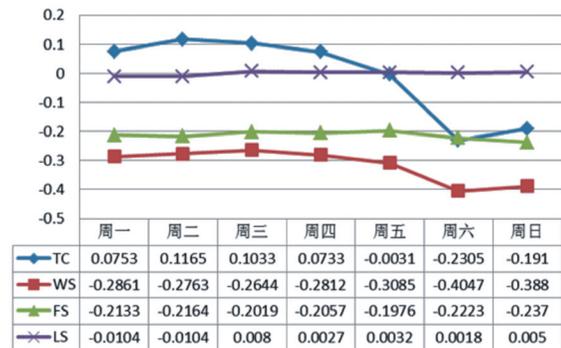


图 3 TC,WS,FS 和 LS 的变化

这一发现与传统研究的结论有所不同。传统的研究认为, 较高的冲突, 将会导致较低的三种满意度。但是本研究的观测发现, 冲突导致较低满意度的情况只存在在工作领域当中, 而其他两个领域中并不明显。

图 4 展示了基于压力的冲突和三种满意度的变化模式。

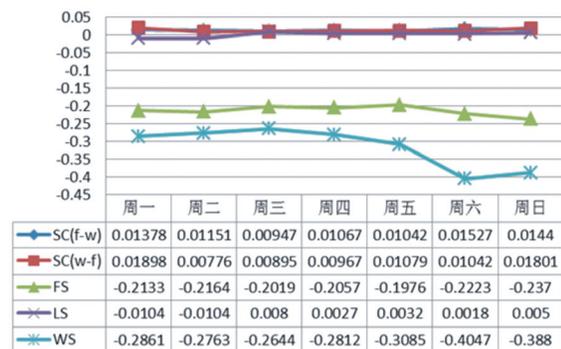


图 4 SC,WS,FS 和 LS 的变化

从图中看到, 在同一坐标系下, 与三种满意度的变化相比, 两种基于压力的冲突都呈现出相对平稳的状态, 周中与周日均没有太大的波动和差别。这与工作满意度和家庭满意度明显缺乏同步。虽然看上去与生活满意度 LS 相近, 但是最终的回归分析显示, 基于压力的冲突与 3 个满意度要素之间均没有明显的相关关系。

可以认为, 基于压力的冲突并不显著影响美国人对三种满意的感受, 即使这种冲突的程度保持一定范围内的稳定, 美国人依然会体会到不同



变化模式的三种满意。这意味着可能存在其他影响因素，能够更大程度上影响三种满意的变化。

综上所述，除了时间的分配仍然影响到美国人对工作的满意情况，其他情形下，工作-家庭冲突与相关的满意之间没有显著的关联关系。这对当下的主流研究结论提出了挑战。

对于三个满意度之间进行的回归分析显示，即使看上去 FS 和 LS 的变化趋同，实际上 WS 和 FS 均没有对 LS 发挥重要作用 (WS 和 LS, -0.2829915 ; FS 和 LS: 0.09994295 , 均达不到 0.05 的值)。也就是说，美国人对生活的满意并不受到家庭满意和生活满意的明显影响，这提出了另一个有趣的研究课题：现在美国人的生活内容都有哪些？他们的价值观是否发生了转变？换言之，是否家庭生活和工作已经不再是他们生活的重心所在？

4 结论与展望

IT 技术、互联网和通信技术的应用模糊了传统的家庭和工作领域之间的界限，人们在任何时间和任何地点都可以从事多领域的活动。这种重叠对领域内的活动产生了各种各样的影响，引起政府、企业和研究者多方面的关注，工作-家庭冲突研究由此成为管理学和心理学领域的热点之一。

然而，众多研究中冲突情形和强度的不同引发了基础性的问题：工作-家庭冲突究竟是一个静态的截面状态，还是一个持续变化的动态过程？如果是后者，冲突与工作满意度、家庭满意度和生活满意度之间的作用关系是否依然存续？

本研究从这两个基础问题出发，在反思传统问卷方法局限的基础上，采用大数据分析的方法，从时序的观点来考察美国工作-家庭冲突的变化模式。结果发现，工作-家庭冲突在星期为单位的时间内有所变化，但是在更长的时间内却呈现相对稳定。这种稳定反映出美国人对 IT 等技术引发的工作-家庭冲突已经适应。分析也发现，当前工作-家庭冲突对三种满意度的影响已经不再明显，而且工作满意和家庭满意不再显著作用于生活满意。研究也表明，大数据分析将带来对传统调查问卷方法的挑战。

本研究还只是本领域内大数据分析方法的初探，今后可以从三个方向继续深入研究，包括：

使用更大数量的数据集，在更长的时间周期内对同样问题进行观测和研究，得到更具普遍性的研究结论；

完善数据分析的公式，改进数据处理的方法，进一步提升研究的可靠性；

研究表明，在不同的管理学领域，数据及其分析和处理具有独特性，可以将大数据分析方法进行更多管理领域的应用研究。

参考文献

- [1] Michel, J. S., L. M. Kotrba, J. K. Mitchelson, M. A. Clark and B. B. Baltes, 2011, "Antecedents of Work-Family Conflict: A Meta-

Analytic Review", *Journal of Organizational Behavior*, 32 (5): 689-725.

- [2] 林忠, 鞠蕾, 陈丽. 工作-家庭冲突研究与中国议题: 视角、内容和设计. *管理世界*, 2013.9, 154-171
- [3] Ellen Ernst Kossek and Cynthia Ozeki (1998). Work-Family Conflict, Policies, and the Job-Life Satisfaction Relationship: A Review and Directions for Organizational Behavior-Human Resources Research. *Journal of Applied Psychology* 1998, Vol. 83, No. 2, 139-149
- [4] Michel, J. S., L. M. Kotrba, J. K. Mitchelson, M. A. Clark and B. B. Baltes, 2011, "Antecedents of Work-Family Conflict: A Meta-Analytic Review", *Journal of Organizational Behavior*, 32(5): 689-725.
- [5] Bagger, J., Li, A., & Gutek, B. 2008. How much do you value your family and does it matter? The joint effects of family identity salience, family-interference-with-work, and gender. *Human Relations*, 61: 187-211.
- [6] Carr, J., Boyar, S., & Gregory, B. 2008. The moderating effect of work-family centrality on work-family conflict, organizational attitudes, and turnover behavior. *Journal of Management*, 34: 244-262.
- [7] Jeffrey H. Greenhaus and Nicholas J. Beutell. Sources of Conflict between Work and Family Roles. *The Academy of Management Review*, Vol. 10, No. 1 (Jan., 1985), pp. 76-88
- [8] Grzywacz, Joseph G., Almeida, David M., and Daniel A. McDonald. 2002. Work-Family Spillover and Daily Reports of Work and Family Stress in the Adult Labor Force. *Family Relations*. 51: 28-36
- [9] Taylor, B. L., R. G. DelCampo and D. M. Blancero, 2009. Work-Family Conflict/Facilitation and the Role of Workplace Supports for U.S. Hispanic Professionals. *Journal of Organizational Behavior*, 30(5): 643-664

Research on Work-Family Conflict Temporal Pattern based on Twitter Data

Liu Yili

Abstract: Work-Family Conflict is one of the most important topic in Organization Behavior and Human Resource Management. Relationships between WFC and satisfaction of work, family and life and its degree are so different in previous research that need more reasonable facts and data to explain. This essay apply big data analysis method to analyzing American Twitter users' temporal pattern. The results show that WFC in America is relatively stable in a long period but fluctuant within a week. Time-based WFC affects work satisfaction obviously while no significant linear correlation between WFC and satisfactions of work, family and life exist. Limitations of survey with sectional data may be the potential reason in previous research.

Key words: work-family conflict; big data analysis; Twitter