High School Principals' Ability to Estimate Work Time

Kennedy O. Ongaga*
Missouri State University, USA

Rebecca Y. Stallings
Missouri State University, USA

Abstract: Time management for educational leaders has remained highly relevant to scholars, policymakers and practitioners. We analyzed survey responses from 98 public high school principals to examine the congruency between average total hours they worked per week against the sum total of the average hours worked per week in each of five distinct categories of leadership tasks. The observed congruence was 0.32, while Cohen's kappa coefficient was 0.10. Female principals tended to underreport, and male principals tended to overreport, total work time. Principals with doctorate degrees exhibited higher congruence than those without, and overreporting was inversely related to highest degree. Principals in charge of large teaching staffs were more likely than their counterparts to be congruent and less likely to overreport total work time. Self-report appears to be an inaccurate method to measure time use among high school principals. If time use is a key component of the quality of principal leadership, more detailed and robust techniques for collecting time use data should be utilized in future studies.

Keywords: *Principals' time use; principalship; congruence of time use measurements; instructional leadership.*

Introduction

The Dimensions of School Leadership Study (DSLS) examined leadership practices and characteristics of public high school principals working under a variety of geographic, demographic, and social contexts in the state of Missouri (Ongaga et al., 2020). There has been a growing body of international research examining the nature and purpose of school principalship and its relationship to school improvement (Andrew et al., 1986; Andrew & Soder, 1987; Day & Sammons, 2014; Gurr, 2017; Gurr et al., 2007; Gurr & Day, 2014; Leithwood et al., 2004; Mulford et al., 2008). The strong link between principal leadership and student achievement has caused educational leadership to be recognized as one of the most influential school-level factors on student achievement (Brockmeier et al., 2013; Fullan, 2001; Hallinger & Heck, 1998; Leithwood et al., 2004; Marzano et al., 2005; Randall & Jones, 1988; Sergiovanni, 2001; Valentine & Prater, 2011). According to Horng and Loeb (2010), effective principals are those who focus on improving student achievement and are "hands-on leaders, engaged with the curriculum and instruction issues, unafraid to work directly with the teachers, and often present in classrooms" (p. 66). However, studies on how principals spend their time (e.g., Grissom et al., 2010; Horng et al., 2010) have shown that they rarely measure up to this ideal. Multiple researchers have found that principals spend minimal amounts of time on instructional leadership activities (Grissom et al., 2015; Horng et al., 2010; May & Supovitz, 2011; May et al., 2012; Murphy, 1990; Murphy et al., 2007). In the review of the literature, we present a variety of methods utilized in educational leadership research to quantify principals' time use, each with its own pros and cons.

Literature Review

Self-Report Time Use Surveys

Self-report surveys have been widely used to solicit information about use of time in different areas of research, largely due to their low cost and low burden for the respondents (Fink, 2017). However, self-report surveys have been criticized for their vulnerability to recall bias, particularly those in which respondents are asked to recall activities or experiences which are not recent (Camburn et al., 2010) or cover a long (Tourangeau et al., 2000) or indefinite (e.g., 'on average') time period. Lavigne et al. (2016) analyzed self-report data from 5950 principals collected during the 2011/12 school year through the U.S. Department of Education's National Center for Education Statistics Schools and

---

* **Corresponding author:**
Kennedy O. Ongaga, Missouri State University, Department of Counseling, Leadership & Special Education (CLSE), USA.
✉ KennedyOngaga@MissouriState.edu

Staffing Survey (2014). Principals were asked to estimate total hours spent per week for their job and the percent of time spent on each of five job-related activity groupings. The researchers converted the five percentages provided into numbers of hours based on the principal's estimated total weekly work hours. While a sample size of this magnitude may mitigate error associated with recall bias, the researchers provide a caveat that findings based on this method "should be interpreted with caution. It is possible that principals either underestimated or overestimated time spent on the job or on certain tasks" (p. 8).

Another methodological issue that often arises with time use surveys is related to cyclical variation over the school year (Camburn et al., 2010). They argued that the way a principal spends their time in the early part of the school year is likely to be quite different from the middle or end of the school year. Hence, a study design that solicits time use information about a single day or week (say, the prior day or week) in order to reduce recall bias is still subject to the effect of cyclical variation. This issue applies as well to a study design that solicits time use information 'on average', because a principal's recall is likely to be influenced by their *recent* experiences.

Camburn et al. (2010) study provides strong evidence against relying on reports or observations based on a single time-point or indefinite time period. They conducted a study in which principals were asked to complete daily logs for seven 5-day periods over a 2-year timeframe. This design permitted them to decompose the variance of their time use measures for six leadership functions using a hierarchical linear model (HLM). The estimated proportions of variance *between principals* were much smaller (.001 to .283) than the proportions of variance *between days* (.717 to .999). The authors reference Weick (1996) and argue that it is reasonable to "expect principals' emphasis on a particular leadership function to vary substantially from day to day. Indeed, the vast majority of the variation in principal's engagement in the six leadership functions was found in day-to-day fluctuation" (p. 721). They further observe:

> Variance estimates from the HLM analyses…clearly indicate that either measuring a single day or asking principals to provide a summary of their practice across a long span of time, such as an entire school year, may misrepresent what principals do because important fluctuations in their work would be obscured. (pp. 730-731)

Some researchers have examined use of time without attempting to estimate hours at all (Augustine et al., 2009; Ghamrawi & Al-Jammal, 2013). Ghamrawi and Al-Jammal asked 60 private-school principals from different regions of Lebanon to rate leadership tasks using a 4-point Likert scale with values 1=*rarely/never*, 2=*occasionally*, 3=*often*, and 4=*always*. Augustine et al. asked 598 principals in 17 school districts in various regions of the U.S. to rate time spent on nine instructional leadership tasks over the past school year using a simple dichotomous scale of 0=*no time or some time* and 1=*a great deal of time*.

Ghamrawi and Al-Jammal (2013) point out a limitation common to all self-report data, regardless of method, stating, "it would be more valid to request the admin team surrounding the principal to complete surveys as well" (p. 64). Grissom and Loeb (2011), who examined self-report effectiveness ratings from 279 principals on leadership skills, concurred:

> While principals' self-ratings reflect an informational advantage in the sense that principals experience themselves performing the tasks, principals are not unbiased observers and may not provide objective assessments. One means of evaluating the validity of the self-assessments is to check them against the ratings of another observer. (p. 1111)

To this end, the researchers administered the same survey instrument to assistant principals in the same school district. They observed that effectiveness ratings from the assistant principals were, on average, lower and more variable than the principals' self-ratings.

Structured Observation

One of the research methods that is known to produce accurate data is structured observation, which is real-time recordings of time spent throughout the workday on individual tasks (Glantz et al., 2019; Bentley et al., 1994; Martinko & Gardner, 1990). Under this method, a trained observer "shadows" the principal during their workday and records, at set time intervals (e.g., 5 minutes), the activity occupying the principal at the moment. The log sheets can either be free-form, in which the observer jots down a brief description of the activity, or grid-based, in which the researchers have devised *a priori* a list of potential activities such that the observer simply checks a relevant box. When free-form log sheets are used, researchers are required to tediously categorize the activity descriptions for all of the participants in order to analyze the data. Even with a grid-based log sheet, the structured observation method can be expensive to implement for a large sample, due to costs related to training and incentivizing the observers. Another disadvantage of structured observations is that a principal might either alter their behavior due to the presence of an observer or might be uncomfortable with the arrangement and thus be reluctant to participate, particularly if the design calls for multiple days of observation.

Camburn and Barnes (2004) conducted a mixed-methods study which included a language arts log with over 150 pre-defined items measuring instruction. The log was filled out at the end of language arts lessons by the observed teachers themselves and two trained observers. They found that agreement between observers was higher than agreement

between observers and teachers. One important interpretation of this finding was that teachers' contextual knowledge, experience, intent, and long-range goals are part of the interactive process, and observers sometimes lacked crucial contextual information that teachers possessed.

In their study to validate structured observations, Mann et al. (1991) made continuous recordings of mother-infant pairs for one hour and used trained observers to record behaviors on log-sheets using different time intervals. They found that both the length of the time interval (i.e., 5 to 60 seconds) and the real-time duration of the behavior affected the mean percent error in the behavior durations estimated from the simulated structured observation data. As length of the time interval increased, mean percent error increased, and as real-time duration of a behavior increased, mean percent error decreased.

Horng et al. (2010) used structured observations with grid-based log sheets to record 43 pre-defined tasks on a single day at five-minute intervals. They postulated that observations should take place on multiple days and at multiple time-points throughout the school year to reduce error due to day-to-day and cyclical variation.

Experience-Sampling

Under experience-sampling (ESM), a principal is prompted electronically (e.g., beeped) at either fixed (e.g., hourly) or random time-points to record their current activity using a self-administered log application listing pre-defined tasks. This method has been shown to have good short- and long-term reliability for assessing frequency and patterning of daily activity and social interaction, among other things (Csikszentmihalyi & Larson, 1987). In the school leadership literature, Spillane et al. (2007) used ESM to estimate principals' work-time in various categories and have demonstrated good validity in comparison to observers' free-form recordings of activities. However, they observed drop-off in compliance during the six consecutive days of recording. Hormuth (1986) reviewed the use of ESM in the personality psychology literature. A post-investigation questionnaire shed some light on non-compliance. Two of the most frequently given reasons for not answering a prompt included: *It is frequently not possible to fill out the questionnaire immediately* (65% yes) and *Filling out the questionnaire disrupts my daily routine* (22% yes). While using the ESM method would not seem to be burdensome time-wise for the principal, it may be viewed as a distraction or inconvenience, leading to low compliance rates. Horng et al. (2010) mention this possibility among comments regarding the ESM method, stating: "The method, however, still suffers from the potential biases inherent in self-reporting. An additional drawback to ESM is that the surveys take time to complete and are thus necessarily limited in their scope so as not to overly disrupt the principal's workday" (p. 2). Additionally, depending on the complexity of the ESM log application, a training session for the participants may add significantly to time and expenses. However, training is necessary to ensure the integrity of data.

Daily Logs

In recent years, educational leadership studies have often utilized a closed-ended time allocation diary, or daily log (e.g., Augustine et al., 2009; Correnti, 2007; Sebastian et al., 2018). Such logs are completed only once per day, and are sometimes referred to as end-of-day logs. Camburn et al. (2010) tested the use of a web-based, closed-ended daily log with 48 principals to estimate time use for nine domains of responsibility. The study design called for the log to be completed at the end of the workday for five consecutive school days at each of seven time-points over a 2-year period. Camburn et al. found good congruence, overall, between the category-specific proportion of time estimated from hierarchical linear modeling (HLM) of the daily log and ESM data (separately). The largest difference in estimated proportion of time was for the category student affairs, whose estimate from the daily log data was 3.45% higher than that from the ESM data. The level 2 variance component of an HLM model estimates between-subject variance. The authors found fairly large discrepancies in between-subject variance estimates from the daily log data and ESM data for the categories *building operations* (daily log variance smaller) and *finances* (daily log variance larger). They observed that:

> ...The differences may reflect limitations in how well the sample of observations obtained for the experience-sampling instrument on a given day represent the full range of activities that occurred on that day. ...Whereas the daily log instrument is a retrospective recall of all activities for a given day, the experience-sampling instrument captures a random sample of activities (p. 722).

Under the assumption that the ESM method is a benchmark, the authors concluded that principals underreported time spent on *building operations* and *finances.* They derived four potential explanations for underreporting: 1) activities were brief and comprised a minor fraction of all activities over the hour (e.g., signing paperwork); 2) activities occurred in a noncontinuous manner (e.g., attempting to write a memo but being frequently disrupted); 3) time bias – activities during the middle of the hour were reported less accurately; and 4) activities were overshadowed by an urgent or dramatic event.

Stylized Questions

Stylized questions such as those used in the DSLS raise concerns about bias and recall error. In a workshop report published by the National Research Council (2000), the participants concluded that there are several reasons why stylized questions can be associated with a significant amount of error. First, activities that are considered socially preferable may be overreported. For example, in the DSLS survey instrument, spending time on *External Relations* (i.e., liaisons with community organizations) may be perceived as a socially desirable activity. Second, it may be difficult for a respondent to conceptualize the timeframe in question, e.g., an "average" workweek. Third, some activities which are being monitored may occur simultaneously. Last, the researcher must anticipate the activities which may occur and define them *a priori* in the data collection instrument. Citing Goldring et al. (2008) and Scott et al. (1990), Horng et al. (2010) note that "recent advances in self-report data collection methods, such as end-of-the-day logs and experience sampling methods (ESM), have reduced some of these potential biases" (p. 492).

Methodology

Research Questions

The DSLS includes a section in which principals are asked to estimate the average number of hours per week spent in total and on tasks that fall into five leadership categories. The inclusion of these items affords us the opportunity to assess the validity of self-report time use by examining the congruence of responses. This study is guided by the following research questions:

1. How strong is the congruence between a principal's estimate of the average number of hours worked per week in total and the sum of the average number of hours worked per week on each of five distinct categories of leadership activities?
2. Does the extent of congruence between these two measures vary by characteristics of the principals or their schools?

Survey Design and Sample Selection

Data for this study primarily comes from a survey that was distributed to non-chartered public high school principals in the state of Missouri. The DSLS form and protocol were reviewed and approved by Missouri State University's Institutional Review Board (IRB). The names, school addresses, and email addresses of all non-chartered public high school principals in the state of Missouri were obtained from a database maintained by the Missouri State Department of Elementary and Secondary Education (DESE) in July 2019. Two modes of completing the survey were utilized simultaneously, namely: a link to an online version of the survey designed using Qualtrics (https://www.qualtrics.com) and a paper version of the survey. As per the instructions, each principal was guided to complete either the online or paper version of the survey by a three-week mailing deadline.

Surveys were mailed to 503 high school principals. Thirty-one surveys were undeliverable by the United States Postal Service. Twenty-three email replies were received stating that the targeted principal had either retired or changed jobs. Twenty-six principals sent out-of-office automatic email responses indicating that they were unavailable until after the survey deadline. One hundred twenty-two surveys were returned (65 paper and 57 online). Based on the number of principals who received the survey before the deadline, the response rate was 29% (122 of 423). Twenty-four of the returned surveys (20%) were dropped from analysis because they were insufficiently complete, leaving a final sample of 98 (see Table 1). The completed paper survey data was entered into an Excel worksheet and the resulting records were combined with those derived from the online responses. Unique identifiers were added to the records in the combined file. The Excel worksheet was imported into SPSS version 24.0 for data cleaning and analysis.

A recent analysis was conducted to examine the representativeness of the 98 schools led by the principals in the final sample relative to all non-chartered public high schools in Missouri. The results confirm that the final sample is representative of all non-chartered public high schools in Missouri in 2018. This analysis is available upon request from the corresponding author.

Table 1. Principals' Personal and School Characteristics (n = 98)

Personal Characteristics	%
Gender	
Female	26
Male	
Age	
<=45	43
><45	57
Highest degree	
Masters	21
Specialist	51
Doctorate	28

Table 1. Continued

Personal Characteristics	%
No. of years served as principal	
<=5	41
6-10	29
11-15	19
>=16	11
No. of years served as assistant principal	
None	41
<=5	36
6-10	17
11-15	6
No. of years served as schoolteacher	
<=5	17
6-10	46
11-15	19
>=16	17
Total no. of years worked in current school	
<=5	35
6-10	25
11-15	14
>=16	26
Total no. of years worked as principal in current school	
<=5	55
6-10	27
11-20	18
School Characteristics	
Location	
Urban	5
Suburban	19
Rural	76
Total no. of enrolled students	
<=100	8
101-500	56
501-1000	14
1001-1500	8
1501-2500	13
Total no. of teaching staff	
<=20	15
21-50	54
51-80	9
81-100	9
>=101	12
Students on free or reduced-price lunch (%)	
<=20	10
21-40	19
41-60	39
61-80	24
81-100	8

Leadership Activity Categories

Principals were asked to estimate the average total number of hours per week they spend working at school (*Total Hours Worked* or *THW*) and on each of five categories of leadership activities (Grissom & Loeb, 2009; Spillane et al., 2007). We labeled the categories/domains as follows: (1) Curriculum and Instructional Leadership (CIL); (2) Internal School Management (ISM); (3) Organization Management (OM); (4) Internal School Relations (ISR); and (5) External Relations (ER) (see Appendix). Based on a review of studies with similar categorization (see Camburn et al., 2010; Grissom & Loeb, 2009; Horng et al., 2010), we believe the domains fairly exhaustively cover the range of principal responsibilities (see Table 2). Horng et al. (2010) and Sebastian et al. (2018) estimated (using different methods), from real-time recordings of activities of the same sample of principals, that between 7.7% and 18.8% of work time in the daily schedule was not utilized for "work" duties. We contend that there is no firm basis on which to dispute that these "transitional" activities should not be randomly distributed across our five categories and that, therefore, any "error" attributable to unrecorded time spent on those activities should be equivalent across the five categories.

Table 2. Comparison of Selected Categorizations of Leadership Activities

Source	Ongaga et al., 2020	Grissom & Loeb (2011)	Horng et al. (2010)	Camburn et al. (2010) and Sebastian et al. (2018)
Study design	Self-reported average hours per week estimates using proscribed time intervals and task categories.	Not a time-focused study. Items rated for effectiveness. EFA used to form the dimensions.	Structured observations in five-minute intervals using proscribed task categories on a single day from 7 am until 6 pm.	Web-based self-administered daily log with hourly prompts for proscribed task categories. Design called for seven 5-day logs over a two-year period.
Grouping title	Principals' Leadership Categories	Dimensions of Principal Skills	Principals' Time Use Categories	Principals' Domains of Responsibility
Grouping taxonomy	Curriculum & Instructional Leadership	Instruction Management	Instructional Program Day-to-Day Instruction	Instructional Leadership Planning & Setting Goals
	Internal School Management	Administration	Administration	Student Affairs
	Organization Management	Organization Management	Organization Management	Building Operations Finances Personnel Issues
	Internal School Relations	Internal Relations	Internal Relations	Parent Relations[1]
	External Relations	External Relations	External Relations	Community Relations[1] School District Functions Professional Growth (personal)
			Other Tasks (includes personal professional development)	Other

[1] *Note*: These authors combine parent and community relations into a single domain labeled *Community or Parent Relations*.

Computation of Time Use Measurements

In the survey instrument, response choices for time use variables were given as intervals of hours. Interval midpoints were substituted to ease comparisons and to allow items to be added together. The interval '<=10' was assumed to mean '0-10'. The open intervals '>40' and '>60' were assumed to have a width equivalent to the closed intervals, i.e., 10 (Hanneman, et al., 2012). The calculated midpoints were rounded down to the nearest integer to ease interpretation. For *Total Hours Worked,* the initial response choices and the final substituted midpoint values are as follows:

Response choices:	<=10	11-20	21-30	31-40	41-50	51-60	>60
Substituted midpoints:	5	15	25	35	45	55	65

For each category of leadership activities, the initial response choices and the final substituted midpoint values are as follows:

Response choices:	<=10	11-20	21-30	31-40	>40
Substituted midpoints:	5	15	25	35	45

These latter substituted values were then utilized to compute a new time use measurement, the sum of time spent on each of the five categories of leadership activities (*Sum of Leadership Activities* or *SLA*).

Measure of Congruency

It is possible that a responding principal could *either* greatly overestimate their average total hours worked per week *or* greatly overestimate their average hours worked per week for one or more category of leadership activities (or even both). We define 'underreport' as an instance when the reported total hours worked per week (THW) is less than the summated total hours worked (SLA). We then devised a simple three-point scale to represent the extent of congruence

between reported (THW) and computed (SLA) total work time with the following definition: '1' represents cases for which there is potential underreporting of average total hours worked (i.e., THW < SLA); '2' represents cases with congruence (i.e., THW = SLA); and '3' represents cases for which there is potential overreporting of average total hours worked (i.e., THW > SLA).

Data Analysis

In the analysis, we utilized Cohen's *kappa* to assess agreement between the reported average total hours worked per week (THW) against the summated average hours worked per week in each of the five leadership activity categories (SLA). Cohen's *kappa* coefficient is a measure of agreement which is valid when two variables are measured using the same response scale and that scale takes on a limited number of distinct values (Chen, 2019).

We also used Somers' *d* (Somers, 1962), a nonparametric measure of association, to assess the association of the three-point measure of congruency (regarded as an ordinal dependent variable) and classes of various personal and school characteristics (regarded as ordinal independent variables). It is worth noting that any dichotomous variable can be treated like an ordinal variable (Agresti, 2010); consequently, it is valid to use Somers' *d* to assess a relationship between an ordinal and a dichotomous variable.

Results

Agreement between Total Hours Worked and Sum of Leadership Activities

The bivariate relationship of the two variables is depicted using a pyramid chart (see Figure 1). The distribution of total work time reported by the principals, *Total Hours Worked*, is symmetric and narrow. The modal response was 51-60 hours/week (midpoint of 55) and the distribution of summated total work time, *Sum of Leadership Activities*, is nearly symmetric and has the same modal response, but is much wider. Basically, it simulates a normal distribution, as anticipated by application of the Central Limit Theorem.
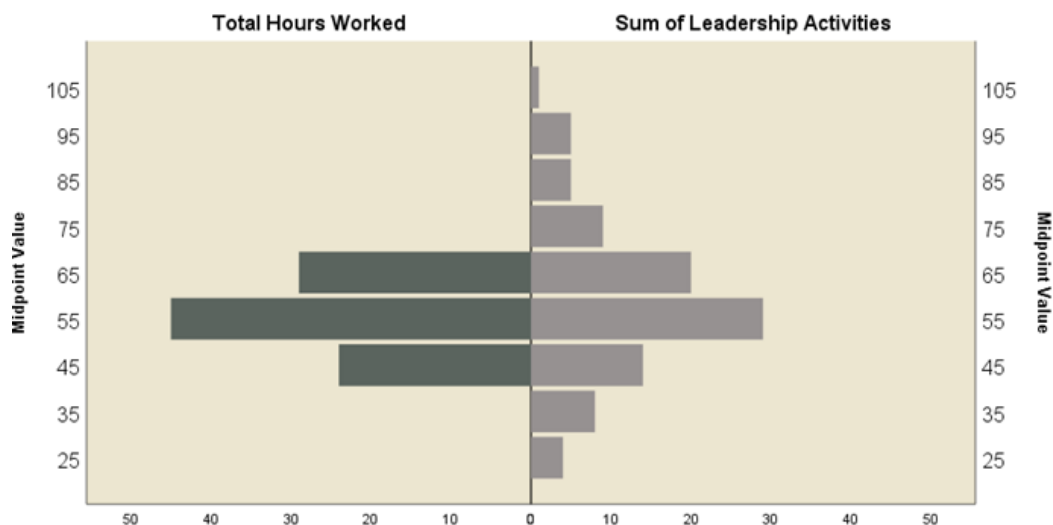

Figure 1. Pyramid chart of Total Hours Worked against Sum of Leadership Activities

Note: From "High School Principals' Leadership Roles and Use of Time," by K. Ongaga, M. Ombonga, K. Finch, and J. Turner, 2020, *Journal of Education and Practice*, *11*(18), p. 8. Copyright 2020 by IISTE.

Total Hours Worked was cross-tabulated by *Sum of Leadership Activities* to examine congruence (see Table 3). Reported total work time and computed total work time were congruent for 32% (30/95) of the principals. Applying the scale developed by Landis and Koch (1977) for Cohen's *kappa* coefficient, this simple observed statistic (0.32) indicates 'fair' agreement. Cohen's *kappa* coefficient for this comparison was much smaller at 0.10, indicating 'slight' agreement, but was statistically significant ($p = 0.05$).

For another 41% (39/95) of principals, their reported total work time was less than their computed total work time. This could be attributable to either overestimating their time spent on one or more leadership activity category or underestimating the magnitude of their average workweek. For the remaining 27% (26/95) of principals, their reported total work time was greater than their computed total work time. This could be attributable to either overestimating the hours in their average workweek or underestimating their time spent on one or more category of leadership activities

Table 3. Total Hours Worked by Sum of Leadership Activities[1]

THW	25	35	45	55	65	75	85	95	105	Total
25	0	0	0	0	0	0	0	0	0	**0**
35	0	0	0	0	0	0	0	0	0	**0**
45	4	1	5	6	5	0	0	1	0	**22**
55	0	5	6	18	8	7	0	1	0	**45**
65	0	2	3	5	7	2	5	3	1	**28**
75	0	0	0	0	0	0	0	0	0	**0**
85	0	0	0	0	0	0	0	0	0	**0**
95	0	0	0	0	0	0	0	0	0	**0**
105	0	0	0	0	0	0	0	0	0	**0**
Total	**4**	**8**	**14**	**29**	**20**	**9**	**5**	**5**	**1**	**95**

Note. [1]Midpoint value assigned to each time interval on both axes. From "High School Principals' Leadership Roles and Use of Time," by K. Ongaga, M. Ombonga, K. Finch, and J. Turner, 2020, *Journal of Education and Practice*, *11*(18), p. 8. Copyright 2020 by IISTE.

Personal and School Characteristics Associated with Congruence

The congruence between reported (THW) and computed (SLA) total work time was further explored by crosstabulating the constructed three-point measure of congruency against key personal and school characteristics of the principals and assessing those relationships using the nonparametric Somers' *d* measure of association. Two personal characteristics demonstrated significant associations with the congruency variable. Female principals were much more likely (56%) than male principals (36%) to 'underreport' total work time (Somers' *d* = -0.238; *p* = 0.050), and much less likely (16%) than male principals (31%) to 'overreport' total work time. The relationship between the congruency variable and principals' highest degree was not uniform in direction. Principals with a specialist degree were much more likely (52%) than those with either a master's degree (29%) or a doctorate (29%) to 'underreport' total work time. On the other end of the congruency scale, an inverse trend was observed for 'overreporting' total work time by highest degree holders, with 38% of master's degree-holders, 26% of specialist degree-holders, and 21% of doctoral degree-holders 'overreporting'. Because the bivariate relationship was not unidirectional, the Somers' *d* test statistic is nonsignificant (Somers' *d* = -0.049; *p* = 0.588). However, the Pearson chi-square test indicates a moderately significant bivariate association (chi-square = 8.426; 4 df; *p* = 0.077).

One school characteristic, size of teaching staff, was found to be moderately significantly associated with the congruency variable (Somers' *d* = -0.195; *p* = 0.082). Principals from schools with a large teaching staff (>50) were more likely than those from schools with a small staff to be 'congruent' (43% vs. 27%, respectively) and much less likely to 'overreport' (11% vs. 34%, respectively).

Discussion

The primary goal of this study was to assess the validity of self-report time use by principals on leadership tasks. Ideally, a time use study of principals would involve real-time recording of time spent on each category of leadership tasks for a specified number of days at multiple time points throughout the school year using either structured observations (Bentley et al., 1994) or experience sampling (Camburn et al., 2010). However, these methods can be burdensome for the principal and may be a deterrent to participation (Hormuth, 1986; Spillane et al., 2007;). In this study, principals were asked to estimate the average number of hours worked per week in total and in each of five leadership categories. While acknowledging the recall error and recall bias inherent in such estimates, we attempted to get a sense of the magnitude of these limitations.

As anticipated, the constructed total work time measurement, *SLA*, has a larger number of classes (i.e., 9) as compared to the recalled total work time measurement, *THW* (i.e., 3*)*. Hence, the Cohen's *kappa* coefficient is based on a non-symmetrical table with three to nine classes. Chen (2019) conducted a simulation exercise to explore factors influencing the value of *kappa*. She concluded that, in general, a *kappa* based on less than six classes should be interpreted cautiously, while *kappa's* 'resiliency' is maximized when it is based on between six and twelve classes. In this study, with a *kappa* of 0.10, we deduce that the congruence between the recalled and constructed work times is only a slight one (Landis & Koch, 1977).

While it is not possible to discern direction (i.e., over or under) from this cross-sectional study, we were able to explore whether congruency was associated with some key personal and school characteristics. Defining potential 'underreport' as reported total work time being *less* than total work time computed by summating times reported for each of the five categories of leadership activities, female principals were much likelier to 'underreport', while male principals were much likelier to 'overreport', total work time. Perhaps male principals are responding to perceived

societal expectations that professional men should work extra-long hours, while professional women need to attend to the needs of their families at the end of the workday.

It is plausible to argue that the demands, accountability measures, and espoused roles of the principals as instructional leaders present some tensions in practice, including use of time (Hallinger et al., 2020). A principal's time spent on a task may not reflect the importance they place on that activity since their responsibilities are numerous and overlapping. Tharp-Taylor et al. (2009) at RAND Corporation surveyed principals in the Pittsburgh Public Schools, asking the amount of time spent on 10 leadership activity categories during a 'typical' week, plus if they felt that the amount of time was appropriate or should be either *more* or *less*. Four of the five categories in which the largest percentage of principals reported spending more than 15 hours in a typical week (student discipline; compliance with Title I or special education requirements; other management issues; communicating with parents and community) were the same as the five categories in which the largest percentage of principals reported that they should be spending *less* time.

Overreporting was inversely related to the principal's highest completed degree, and principals with a doctorate were the most likely to have congruence between the total time measures. Perhaps the additional years of formal training results in better time management skills and attention to schedule planning. Similarly, principals from large schools were more likely to have congruence and much less likely to 'overreport'. Principals from large schools may be more attuned to their time allocation due to juggling larger staffs and student bodies and more curricular and non-curricular activities.

The findings of this study reflect the comment by Horng et al. (2010, p. 492) that self-report surveys allow for large samples, but often sacrifice depth and, perhaps, accuracy. Studies utilizing self-report surveys (e.g., Lavigne et al., 2016; Ghamrawi & Al-Jammal, 2013) are likely to be susceptible to self-reporting and memory biases.

Conclusion

Self-report surveys appear to be an inaccurate method to measure the use of time. Additionally, the findings suggest that under- or overreporting may not be random. Studies of principals and principalship in which use of time is a key component of the study objectives should include more detailed, albeit more time-consuming and possibly costly, techniques for collecting time use data in order to increase accuracy and ease interpretation of the findings. Further studies are needed to examine principals' work patterns more deeply in order to understand if principals allocate time differently based on different expectations and pressures.

Recommendations

Principals were asked to estimate the number of hours per week spent on a list of leadership activities, which were clustered into five distinct categories. While changes in the principals' roles have brought instruction more to the center of their work, principals need district-level support that will release more time for them to be effective instructional leaders. More research is needed with larger sample sizes that focus on how principals spend their time in different policy and social contexts, including large vs. small schools, urban vs. rural schools, schools with homogenous vs. diverse student populations, and high vs. low poverty schools. More robust data collection strategies need to be utilized to overcome inherent potential biases.

Limitations

While data from the DSLS provide a valuable opportunity to examine principals' ability to estimate work time, there are several limitations. First, we are limited by the fact that only one data collection method, surveys, was used. This is likely to sacrifice depth and, perhaps, accuracy. Second, our sample size is too limited to generalize the findings to most contexts. Finally, principals' use of time in instructional leadership, specifically, needs to be measured against what states and policy makers consider to be of priority for school effectiveness. Qualitative data on principals' voices regarding what makes school leaders effective will embellish such studies.

References

Agresti, A. (2010). *Analysis of ordinal categorical data* (2nd ed.). John Wiley & Sons.

Andrews, R. L., Soder, R., & Jacoby, D. (1986, April 20). *Principal roles, other in-school variables, and academic achievement by ethnicity and SES* [Paper presentation]. American Educational Research Association, San Francisco, CA, United States.

Andrews, R. L., & Soder, R. (1987). Principal leadership and student achievement. *Educational Leadership 44*(6), 9–11.

Augustine, C. H., Gonzalez, G., Ikemoto, G. S., Russell, J., Zellman, G. L., Constant, L., Armstrong, J., & Dembosky, J. W. (2009). *Improving school leadership: The promise of cohesive leadership systems*. RAND Corporation. https://www.rand.org/pubs/monographs/MG885.html

Bentley, M. E., Stallings, R., & Gittelsohn. J. (1994). The structured observation technique for the study of health behavior. In S. Cairncross & V. Kochar (Eds.), *Studying Hygiene Behavior: Methods: Issues, and Experiences.* Sage.

Brewer, D. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review, 12*(4), 281–292. https://doi.org/10.1016/0272-7757(93)90062-L

Brockmeier, L. L., Starr, G., Green, R., Pate, J. L., & Leech, D. W. (2013). Principal and school-level effects on elementary school student achievement. *NCPEA International Journal of Education Leadership Preparation, 8*(1), 49–61.

Camburn, E. & Barnes, C.A. (2004). Assessing the validity of a language arts instruction log through triangulation. *The Elementary School Journal*, *105*(1), 49-73.

Camburn, E.M., Spillane, J.P., & Sebastian, J. (2010). Assessing the utility of a daily log for measuring principal leadership practice. *Educational Administration Quarterly*, *46*(5), 707-737.

Chen, Y. S. (2019, July 5). *Interpretation of Kappa Values: Evaluate the agreement level with condition*. Towards Data Science. https://towardsdatascience.com/interpretation-of-*kappa*-values-2acd1ca7b18f

Correnti, R. (2007). An empirical investigation of professional development effects on literacy instruction using daily logs. *Educational Evaluation and Policy Analysis, 29*(4) 262-295. https://doi.org/10.3102/0162373707309074

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, *175*, 526-536. https://doi.org/10.1097/00005053-198709000-00004

Day, C., Harris, A., Hadfield, M., Tolley, H., & Beresford, J. (2000). *Leading schools in times of change*. Open University Press.

Day, C., & Sammons, P. (2014). *Successful school leadership*. Education Development Trust. https://www.educationdevelopmenttrust.com/our-research-and-insights/research/successful-school-leadership.

Erickson, D. A., & Reller, T. L. (Eds.). (1978). *The principal in metropolitan schools*. McCutchan Publishing Corporation.

Fink, A. (2017). *How to conduct surveys: A step-by-step guide* (6th ed.). Sage.

Fullan, M. (2001). *Leading in a culture of change.* Jossey-Bass.

Ghamrawi, N., & Al-Jammal, K. (2013). Time flies – A statement that best applies to school principals. *British Journal of Education*, *1*(1), 52-66.

Glantz, A., Ormon, K., & Sandstrom, B. (2019). How do we measure time? – An observational study measuring the task time distribution of nurses in psychiatric care. *BMC Nursing 18*(67), 1-8. https://doi.org/10.1186/s12912-019-0386-3

Goldring, E., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration*, *46*(3), 332-352. https://doi.org/10.1108/09578230810869275

Grissom, J. A., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal, 48*(5), 1091-1123.

Grissom, J.A., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher*, *42*(8), 433-444.

Grissom, J., Loeb, S., & Mitani, H. (2015). Principal time management skills: Explaining patterns in principals' time use, job stress, and perceived effectiveness. *Journal of Educational Administration*, *53*(6), 773-793.

Gurr, D. (2017). A model of successful leadership from the International Successful School Principalship Project. In K. Leithwood., J. Sun & K. Pollock. (Eds.), *How school leaders contribute to student success: Studies in Educational Leadership*. Springer.

Gurr, D. (2014). Successful school leadership across contexts and cultures. *Leading and Managing, 20*(2), 75–88.

Gurr, D., & Day, C. (2014). Thinking about leading schools. In Day, C., & Gurr, D. (Eds.), *Leading schools successfully: Stories from the field* (pp. 194-208). Routledge.

Gurr, D., Drysdale, L., & Mulford, B. (2007). Models of successful principal leadership. *School Leadership & Management, 26*(4), 371-395. https://doi.org/10.1080/13632430600886921

Hallinger, P., Gümüş, S., & Bellibaş, M.Ş. (2020). Are principals' instructional leaders yet? A science map of the knowledge base on instructional leadership, 1940–2018. *Scientometrics, 122*, 1629-1650. https://doi.org/10.1007/s11192-020-03360-5

Hallinger, P., & Heck, R.H. (1998). Exploring the school principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement*, *9*(2), 157-191.

Hanneman, R.A., Kposowa, A.J., & Riddle, M.D. (2012). *Basic Statistics for Social Research* (1st ed.). John Wiley & Sons.

Hormuth, S. (1986). The sampling of experiences *in situ*. *Journal of Personality*, *54*(1), 262-293.

Horng, E., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education, 116*(4), 491-523.

Horng, E.L., & Loeb, S. (2010). New thinking about instructional leadership. *Phi Delta Kappan*, *92*(3), 66-69. https://doi.org/10.2307/25753685

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lavigne, H.J., Shakman, K., Zweig, J., & Greller, S.L. (2016). *Principals' time, tasks, and professional development: An analysis of Schools and Staffing Survey data* (REL 2017-201). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http://ies.ed.gov/ncee/edlabs

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*: The Wallace Foundation. https://www.wallacefoundation.org/knowledge-center/Documents/How-Leadership-Influences-Student-Learning-Executive-Summary.pdf

Mann, J., Have, T.T., Plunkett, J.W., & Meisels, S. J. (1991). Time sampling: A methodological critique. *Child Development*, *62*(2), 227-241. https://doi.org/10.2307/1130999

Marzano, R.J., Waters. T., & McNulty, B. A. (2005). *School leadership that works: From research to practice.* Association for Supervision and Curriculum Development. Mid-continent Research for Education and Learning.

May, H., & Supovitz, J. A. (2011). The scope of principal efforts to improve instruction. *Educational Administration Quarterly*, *47*(2), 332–352. https://doi.org/10.1177/0013161X10383411

May, H., Huff, J., & Goldring, E. B. (2012). A longitudinal study of principals' activities and school performance. *School Effectiveness and School Improvement*, *23*(4), 417–439. https://doi.org/10.1080/09243453.2012.678866

Martinko, M. J., & Gardner, W. L. (1990. Structured observation of managerial work: A replication and synthesis. *Journal of Management Studies 27*(3), 329–57.

Marzano, R., Waters, T., & McNulty, B. (2005). *School leadership that works: From research to results.* Association for Supervision and Curriculum Development (ASCD).

Mulford, B., Kendall, D., Ewington, J., Edmunds, B., Kendall, L., & Silins, H. (2008). Successful principals of high-performance schools in high-poverty communities. *Journal of Educational Administration 46*(4), 461-480. https://doi.org/10.1108/09578230810882009

Murphy, J. F., Elliott, S. N., Goldring, E., & Porter, A. C. (2007). Leadership for learning: A research-based model and taxonomy of behaviors. *School Leadership and Management*, *27*(2), 179–201. https://doi.org/10.1080/13632430701237420

Murphy, J. F. (1990). Principal instructional leadership. In L. L. Lotto & P. W. Thurston (Eds.), *Advances in educational administration: Changing perspectives on the school* (pp. 163–200). JAI Press.

Ongaga, K., Ombonga, M., Finch, K., & Turner, J. (2020). High school principals' leadership roles and use of time. *Journal of Education and Practice*, *11*(18), 1-15. https://doi.org/10.7176/JEP/11-18-01

Randall, E. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review, 7*(3), 291–99.

Scott, C., Ahadi, S., & Krug, S. E. (1990). *An experience sampling approach to the study of principal instructional leadership II: A comparison of activities and beliefs as bases for understanding effective school leadership*. The National Center for School Leadership. https://files.eric.ed.gov/fulltext/ED327943.pdf

Sebastian, J., Camburn, E.M., & Spillane, J.P. (2018). Portraits of principal practice: Time allocation and school principal work. *Education Administration Quarterly*, *54*(1), 47-84.

Sergiovanni, T.J. (2001). *Leadership: What's in it for schools?* Psychology Press.

Somers, R. H. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review, 27*(6), 799-811. https://doi.org/10.2307/2090408

Spillane, J.P., Camburn, E., & Pareja, A. (2007). Taking a distributed perspective to the school principal's workday. *Leadership and Policy in Schools*, 6(1), 103-125. https://doi.org/10.1080/15700760601091200

Tharp-Taylor, S., Nelson, C. A., Hamilton, L.S., & Yuan, K. (2009). *Pittsburgh Public Schools' Excellence for All: Year 2 Evaluation*. RAND Corporation. https://www.rand.org/pubs/documented_briefings/DB575-1.html

Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

U.S. Department of Education. (2014). *Schools and Staffing Survey (SASS)* [Data set]. http://nces.ed.gov/surveys/sass/index.asp

Valentine, J. W., & Prater, M. (2011). Instructional, transformational, and managerial leadership and student achievement: High school principals make a difference. *NASSP Bulletin, 95*(1), 5-30. https://doi.org/10.1177/0192636511404062

Weick, K. E (1996). *Sensemaking in organizations.* Sage.

APPENDIX

The following is a reproduction of the time use section of the Dimensions of School Leadership Survey instrument.

Principal's Work and Time Use

In this section, we are interested in how you spend and allocate your time to various leadership categories, including curriculum and instructional leadership, school management, organization management, internal and external school relations. Please estimate the time you spend on the following categories every week.

14. On average, how many total hours per <u>week</u> do you work at school?

<10	41-50
11-20	51-60
21-30	>61
31-40	

15. On average, how many hours/week do you spend on <u>Curriculum and Instructional Leadership</u> (e.g., formally evaluating teachers and providing instructional feedback, implementing required professional development, using data to inform instruction, using assessment results for program evaluation and development, etc.)?

<10	31-40
11-20	>40
21-30	

16. On average, how many hours/week do you spend on <u>Internal School Management</u> (e.g., managing student services, managing schedules, etc.)?

<10	31-40
11-20	>40
21-30	

17. On average, how many hours/week do you spend on <u>Organization Management</u> (e.g., managing personnel, budgets, facilities, etc.)?

<10	31-40
11-20	>40
21-30	

18. On average, how many hours/week do you spend on <u>Internal School Relations</u> (e.g., counseling students and/or parents, attending school activities, communicating with parents, etc.)?

<10	31-40
11-20	>40
21-30	

19. On average, how many hours/week do you spend on <u>External Relations</u> (e.g., working with local community members or organizations, utilizing district office meetings, etc.)?

<10	31-40
11-20	>40
21-30	