MEASUREMENT OF MULTITASKING WITH FOCUS SHIFT ANALYSIS

Completed Research Paper

Raquel Benbunan-Fich
Baruch College, City University of New York
One Bernard Baruch Way, Box B11-220, New York, NY 10010
rbfich@baruch.cuny.edu

Abstract

Perhaps one of the most prevalent, yet least understood, forms of modern behavior is human multitasking aided by technology. Multitasking is both difficult to define and to measure. Based on Activity Theory and the concept of focus shifts, this study defines multitasking as shifts that occur when the flow of work is interrupted and tasks are interchanged during a session. The analysis of focus shift diversity, with respect to the tasks that receive attention, enables researchers to measure multitasking. With this approach, several indices of diversity are examined in order to provide recommendations about alternative measurement choices. This methodological study advances our understanding of the possibilities and limitations of using diversity indices for measuring multitasking.

Keywords: Template, formats, instructions, length, conference publications

Introduction

Human multitasking is one of the most common but least understood forms of modern behavior. Recent academic literature and countless articles in the popular press provide accounts of humans performing apparently simultaneous tasks with the help of modern technology. This growing literature suggests that multitasking can be researched from a contextual, a cognitive, or a technological perspective. The contextual viewpoint determines whether engaging in multitasking is appropriate (McWilliams 2006). For example, the use of laptops in lectures to manage unrelated tasks, like handling email, is contextually inappropriate (Adams 2006). The cognitive angle investigates whether some people are more adept at multitasking than others due to their mental abilities (Watson and Strayer 2010) or their age (Wallis 2006). The technological perspective addresses the pervasiveness of multitasking as a behavior that takes place anytime anywhere, aided by a variety of devices (Benbunan-Fich and Truman 2009). The latter is the perspective adopted in this paper.

Modern technology has removed temporal and geographic barriers that previously constrained activities to predetermined times and places. As a result, contemporary human activities have been reorganized. This spatial and temporal rearrangement, fueled by technology, is characterized by the blurring of the lines between activities that used to be well-separated (e.g. work, family time, entertainment, or social interactions). The lack of demarcation has created an environment where people work in a fragmented way, integrating components of different tasks whenever possible. Because most modern tasks are mediated by technology, it is possible to attend to multiple ongoing activities in a span of time (or session). During a specific period, some pursuits are abandoned midway, for various reasons to attend to other tasks. This task layering provides a glimpse into the complexity of modern work, but it does not offer a complete account of multitasking activity.

Multitasking behavior occurs when tasks currently underway are temporarily abandoned to undertake or to continue other tasks. The temporal arrangement of task segments leads to a zigzagged performance when components of several tasks are combined within the same time frame. This task segmentation and multiple task integration could be triggered by purposeful organization of one's work or by unplanned reactions to external interruptions such as email notifications. Regardless of the cause, there is *a shift in the focus of attention* when tasks are swapped (Waller 2007). We propose that a quantitative analysis of focus shifts provides the foundation for measuring multitasking (Benbunan-Fich 2011b).

Drawing from Activity Theory, this paper revisits the concept of focus shifts originally introduced by Bødker (1996) and applies it to the context of multitasking behavior. The main premise is that by measuring the diversity of focus shifts, one can obtain a quantitative indicator of multitasking behavior. To this end, this paper proposes and analyzes three measures of diversity (richness, Shannon's H and Simpson's diversity index) and outlines a set of principles for the selection of the most suitable measure. The recommended index (Shannon's H) is empirically validated to assess its discriminant ability. In light of this approach, this study aims to offer groundbreaking methodological contributions stemming from the combination of Activity Theory with indices traditionally used in natural sciences to quantify multitasking in a systematic way.

Activity Theory

Activity Theory (AT) provides the foundation to understand the hierarchical nature of human activities and the mediating role of tools in their performance. This meta-theory integrates motivation, cognition, and behavior in the context of human practices (Bedny and Karwowski 2007; Kaptelinin and Nardi 2006). It emphasizes human intentionality while recognizing that an activity is a combination of pre-specified and situated components (Kaptelinin and Nardi 2006). Planned actions and emergent reactions jointly determine the nature of human activities.

The main tenets of AT are human intention, hierarchical conceptualization of activity, and tool mediation. *Human intention* refers to the unity of consciousness and activity when an individual interacts with other people or artifacts. *Hierarchical conceptualization* consists of defining each activity at the top of a hierarchy that includes nested actions, each of which in turn involves subordinated operations. *Tool* *mediation* indicates that humans carry out specific activities through artifacts or tools that connect a person not only with the world of objects, but also with other people (Benbunan-Fich et al. 2011a).

The two main streams of Activity Theory are the Scandinavian school (Kuuti, 1991) and the systemic-structural approach developed by Bedny and colleagues (Bedny and Karwowski 2007). The Scandinavian approach emphasizes the principle of object-orientation whereby an activity is a unit of subject-object interaction (Kaptelinin and Nardi 2006). In contrast, the systemic-structural approach is more focused on conscious goals as drivers of activity (Bedny and Karwowski 2007). The primary object of study in this approach is human activity during task performance. An activity is defined as a "goal directed system in which cognition, behavior and motivation are integrated and organized by goals and the mechanisms of self-regulation" (Bedny et al. 2000, p. 168). Tasks are therefore the means to organize actions (or sub-tasks) around a specific goal. Since this approach explicitly recognizes the existence of tasks as a middle layer between activity and actions, and it is particularly focused on the link between tasks and goals, the systemic-structural perspective provides an appropriate theoretical basis to explain multitasking behavior.

Bedny et al. (2000) and more recently Benbunan-Fich et al. (2011a) have extended the conceptual foundation of Activity Theory to a multiple task context. In essence, this extension acknowledges that tasks are vehicles to organize actions around goals, and that task components can be fragmented and carried out in specific time periods. An activity is regarded as a self-regulative or adaptive system that actively interacts with the situation (Bedny and Karwowski 2007) and multiple activities co-exist in the same time period.

Focus Shifts

Drawing from the original hierarchical conceptualization of activities in AT, Bødker (1996) defines focus shifts as a change in actions or operations caused by breakdowns or by deliberate shifts. More specifically, "a *focus shift* occurs when work is interrupted to focus on the tool at hand" (p. 150). Although the concept was presented in the context of a single task, we expand this definition to a multiple task context as follows: focus shifts occur either when the technology changes, but the underlying task is the same (as in Bødker's original conceptualization); or when the technology changes and the underlying task changes, as well. By explicitly acknowledging whether there is a change in the underlying task, two types of focus shifts can be distinguished, namely: *technology*-focus-shifts and *task*-focus-shifts. In the former, there is a tool adjustment to continue the performance of the same task, while in the latter, the tool is altered to carry out or continue another task. Multitasking occurs only when there is a *task*-focus-shift.

In computer-mediated work, task changes are typically accompanied by technology switches but not all technology switches are indicative of multitasking. For example, an individual who is using data from a spreadsheet to write a report may be constantly switching between two applications (word processor and spreadsheet program) to check, integrate, and eventually import data into the document. These types of shifts are considered technology-focus-shifts because the tool is being adjusted to facilitate the performance of the same task (i.e. writing the report). When the main task does not change, technology focus shifts are not representative of multitasking behavior.

Alternatively, an individual could be writing a report with information held in his memory (without consulting any other sources). While crafting the report, he could be intermittently switching to handle unrelated email or to check websites unconnected with the content of the report. These focus shifts indicate the temporary abandonment and subsequent pursuit of different goals, and are thus indicative of multitasking. Focus shifts that represent task changes are germane for the study of multitasking behavior. For this reason, the remainder of this paper considers only task focus shifts, which for brevity are called simply focus shifts.

We propose that focus shifts provide good indicators for understanding and measuring multitasking behavior. In the next section, we explain how to quantify them and how different indices of diversity applied to focus shifts analysis yield different types of multitasking measures.

Quantification of Focus Shifts

Multitasking behavior results from the decision to collocate components from multiple tasks in the same period of time for their execution. There are three generic strategies for the temporal organization of multiple independent tasks in a specific timeframe: sequential, interleaved, and parallel (Bluedorn et al. 1992). In the *sequential* mode, multiple tasks are carried out in a specific interval but only one task is executed at a time, from beginning to end. In the parallel mode, all tasks are attended at the same time. In the *interleaved* mode, tasks underway are temporarily suspended to perform other tasks and previously abandoned tasks are eventually resumed (Benbunan-Fich et al. 2011a). A special case of interleaving occurs when an ongoing task is suspended to perform another one in its entirety such that the second task is completely *embedded* into the first. This particular case has implications for focus shift quantification, and it is considered separately. Each of these approaches is further elaborated below.

Parallel performance implies that humans are dividing their attention among simultaneous tasks. This mode is not suitable for focus shift quantification because such shifts occur at the cognitive level and are not externally visible. An illustration of this situation occurs when a person writes a paper while listening to music (Waller 2007). Since those two activities happen at seemingly the same time and there are no visible focus shifts that could be quantified, this case is excluded from further consideration.

The three remaining focus shift prototypes – sequential, interleaved and embedded – can be represented in a bi-dimensional graph with *Tasks* on the vertical axis and *Time* on the horizontal axis. The vertical axis shows different tasks with independent goals, each identified with TG (for Task-Goal) and a number. The horizontal axis shows a specific period of time or session, from start to end.

In the *sequential* prototype (Figure 1), multiple tasks are executed in succession, such that each task starts after completing the previous one. The vertical markers indicate the beginning or conclusion of a task depending upon its location to the left or right of the task segment. The t_1 and t_2 indicators just mark transitions between tasks. Although multiple tasks are performed in the same time period, there is no segmentation of ongoing tasks. As a result, the total number of focus shifts is zero because no task was stopped before its conclusion to attend to another. Therefore, this pattern is not representative of multitasking behavior.

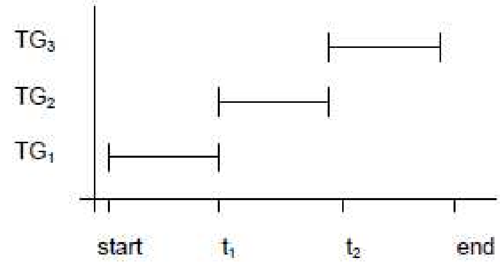

Figure 1. Sequential Prototype

The *interleaved* approach (Figure 2) consists of interweaving tasks, or task components, by diverting attention or shifting the focus from one task to another and eventually resuming previously abandoned tasks. In the horizontal axis of this figure, each change of task is marked with the letter "f" and the corresponding subscript. The diamond symbols at the beginning or at the end of each task segment indicate interruption or resumption. Focus shifts occur at each f_i when there is a breakage in the flow of work, and consequently a focus shift. Figure 2 shows three tasks and six focus shifts, marked with f_1-f_6. There are only six because neither the beginning nor the end of the session count as focus shifts.

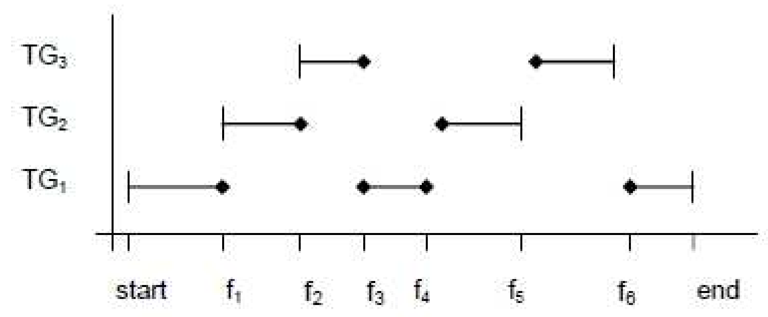

Figure 2. Interleaved Prototype

Embedding is a different strategy for the temporal organization of multiple tasks that results in fewer focus shifts (Figure 3). It occurs when the performance of one task is completely inserted into another. Figure 3 shows three tasks and four focus shifts. The embedded task (TG_3) is not fragmented, and therefore the total number of focus shifts is lower than in the case of total interleaving.

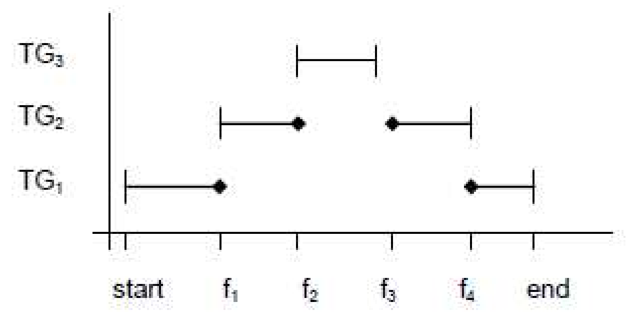

Figure 3. Embedded Prototype

The two multitasking prototypes illustrated in Figures 2 and 3 can be expanded by increasing the number of tasks, by changing the number of focus shifts, and/or by altering the fragmentation and interspersion of task segments. Regardless of the specific pattern of reallocation of attention, *given the same number of tasks, a higher count of focus shifts will indicate a higher level of multitasking*. Said differently, given the same three tasks, a user working according to the pattern described in Figure 2 will be multitasking more than another person working according to Figure 3.

For a more complex illustration, we will use the case of an individual who is handling five independent tasks (TG_1-TG_5) during the period defined by the start and end of his/her work session (See Figure 4). Tasks are numbered by the order in which they were first undertaken. During the session, the person experienced a total of fifteen focus shifts (f_1-f_{15}). For concise notation, these shifts can be counted with respect to the task that receives attention in each case (TG_1 to TG_5). With this notation, the structure of a session is represented with a string showing the number of total shifts per task (e.g. 5-3-1-4-2).

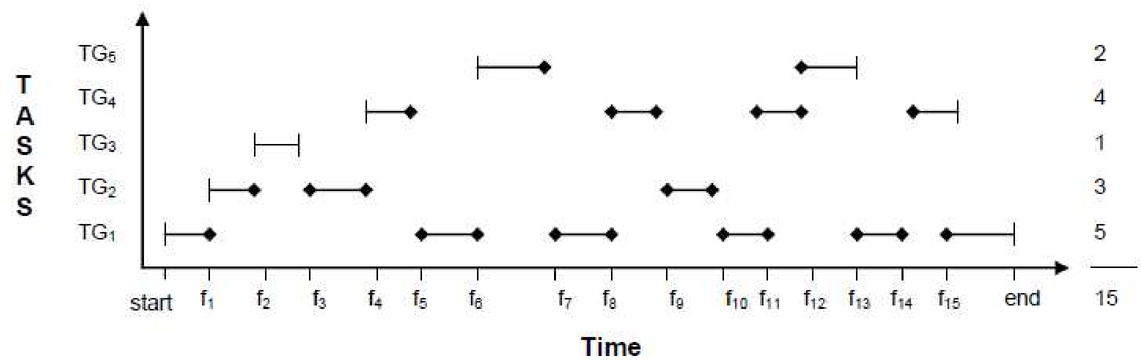

Figure 4. Distribution of Focus Shifts

Figure 4 illustrates several issues associated with focus shift analysis. First, because it consists of a single segment, TG3 is an example of an embedded task. Second, TG2 was left unfinished. In this particular case, task completion or incompletion does not affect focus shift quantification because TG2 was abandoned to continue another task (TG1) that was previously started (i.e. ongoing task). However, when a task is left unfinished (and never attended again during the session) to start a brand new task, this move should be considered a transition and not a focus shift. Third, the intervals representing the times when each task was performed have unequal length. This unevenness in the segments represents more realistic situations in which task fragments are carried out in periods of different duration. Fourth, the aim of this particular scenario is to provide an example with many tasks (5) and a more complex pattern of focus shifts (15). These numbers were selected to be consistent with the example presented in Benbunan-Fich (2011b). In general, the ability to describe a session numerically and graphically using this approach provides an important tool for understanding and quantifying multitasking. The numeric representation of focus shifts is used as the foundation to discuss potential indicators of multitasking behavior in the next section.

Indices of Diversity

In biology, a diversity index is a mathematical measure of variety of the species present in a community or group. Diversity is defined as a function of the proportional distribution of the population members in a set of mutually exclusive and unordered categories. Each population element can be classified into one category. When all population members belong to the same category, diversity has its minimum value of zero. In contrast, diversity is maximized when each category is equally represented in the population (Baczkowski et al. 1997, Desroschers and Anand 2004).

Diversity can be quantified in different ways depending on how the richness and abundance of species are taken into account. Richness indicates the number of species present in a particular group, while abundance compares the proportions of each class present in the selected community or group. Evenness occurs when the relative abundance of the different species is similar. A group has high species diversity if many species are present in nearly equal proportions (evenness). In contrast, if a group has only a few species or if only a few species are very abundant, then species diversity is low (Baczkowski et al. 1997).

The application of these concepts to multitasking behavior measurement requires a specification of “species” and “group.” We define *species as the tasks* that are performed so that each focus shift can be classified with respect to the task that receives attention. We define *group as the total number of focus shifts in a session*. With this mapping, multitasking can be measured with any diversity index. The basic assumption for such an index is that a session dominated by focus shifts in one or two tasks is less diverse than one in which all the tasks receive an equivalent number of focus shifts. Ideally, a suitable index should capture information about richness and abundance of focus shifts in a work session.

The three basic and most widely used indices of diversity are richness, Shannon's H, and Simpson's index of Diversity or SID (Desroschers and Anand 2004). While richness is easy to determine, the other two require the calculation of the proportion of focus shifts (p_i) for each task *i* relative to the total number of all focus shifts (N) in the session.

Before introducing the indices, it is necessary to clarify the notation that will be used. Given a population of N elements from *s* different species (or tasks), such that n_i is the number observed from the i^{th} species, then the relative abundance of species *i* is calculated as: $p_i = n_i / N$, with i=1, 2, 3, ... , s. According to Pielou (1975), there are two main characteristics that a good diversity index should exhibit. First, for a given number of species the index should reach its maximum when the p_i's are equal. Second, if the p_i's are equal for two groups, the index should be an increasing function of s (Barczkowski et al. 1997). Pielou's properties require that for a given level of richness (*s*), the index increases as evenness increases, and for a given level of evenness, the value of the index is greater as richness increases.

Richness

The simplest indicator of diversity is the number of species or tasks active during a session. Generally called *species richness* – or task richness in our case – this indicator gives insufficient information about the relative abundance of focus shifts to each task. In the measurement of multitasking, task richness by itself may be misleading because it may indicate that multitasking is present in cases where focus shifts are zero (as in the prototype shown in Figure 1).

For a more precise description of multitasking activity, since the population frame is determined by a session, it is necessary to provide an indication of the size of the population, or total number of focus shifts observed in the session. These two parameters – richness (or number of tasks) and size (or number of focus shifts) – are important but insufficient because they do not provide information about the distribution of focus shifts during a session. The next two indices of diversity provide information about the dispersion of focus shifts.

Shannon Index (H)

For Shannon's index (H), each p_i is multiplied by the natural logarithm of each proportion ($\ln(p_i)$). The resulting product is summed across tasks, and multiplied by -1:

$$\mathrm{H} = -\sum_{i=1}^{s} p_i \ln(p_i)$$

By convention, 0*ln(0) is defined as zero. It can be shown mathematically that this index fulfills the two properties proposed by Pielou (1975). Given its formulation, H captures both abundance of focus shifts and evenness of their distribution across tasks. This index ranges from 0 (no diversity) to ln(s). If H is divided by its maximum possible value (ln(s)), the resulting index is bounded between 0 (no diversity) and 1 (maximum evenness). This normalized measure (H*) is sometimes called Shannon's equitability index. The application of this formula to the session illustrated in Figure 4 yields the following result: H= 1.49 and H*=H /ln(5) = 0.93

Simpson Index of Diversity (SID)

Simpson's diversity index is based on the calculation of D, which is obtained as the sum of the squared proportions of focus shifts (p_i) for all tasks. In contrast to Shannon's H, D varies inversely with the equitability of the different proportions. In fact, the calculation of D was originally introduced as a measure of dominance or concentration. Over time, several authors have proposed modifications, such as using the reciprocal (1/D) or subtracting it from one (1-D), to transform it into a suitable diversity index (Barczkowski et al. 1997). Because of its normalized properties, we adopt the latter formulation and calculate Simpson's index of diversity (SID) as follows:

$$SID = 1 - D = 1 - \sum_{i=1}^{s} p_i^{2}$$

Although the original sum of squared proportions (D) does not satisfy the properties outlined by Pielou (1975), the formulation of SID as (1-D) does. Its value ranges between 0 and 1, and higher values are associated with greater diversity. For the session reported in Figure 4, SID is 0.76.

Generalization

These three measures of diversity – richness, Shannon's H and SID – are moments of Rényi's (1960) generalized entropy.

$$H_{\alpha} = \frac{1}{1-\alpha} - \ln(\sum_{i=1}^{s} p_i^{2})$$

Richness can be calculated from this formula using $\alpha = 0$, Shannon's H can be obtained with α approaching 1 and SID can be calculated with $\alpha = 2$ (Desroschers and Anand 2004). The values of α are used to identify and label diversity indices. An index computed from the squared proportions or frequencies, such as SID, is called "order 2." If the proportions are used but not squared, as in Shannon H, the index is said to be of "order 1." By extension, an index that is a function of the zero power of the proportions is "order 0." This is the case of richness. The idea of using entropy to measure multitasking activity was first introduced by Benbunan-Fich (2011b).

Evaluation of Indices

Although the indices are equivalent to a certain extent, the decision of which one to use should rest upon well-defined criteria. These considerations enable researchers to understand the possibilities and limitations of each choice. To this end, our critical examination of the indices will address the underlying assumptions required for their application, as well as their most salient properties. This analysis will be followed by the empirical validation of the most promising measure. Based on this evaluation, a set of recommendations is provided.

Aside from access to appropriate data sources, the indices of diversity presented in this study rely on three basic assumptions. First, when applied to the context of multitasking measurement, the session must be well-defined with a clear beginning and end. Second, focus shifts must be identifiable and amenable to classification in mutually exclusive tasks categories. Third, no prior assumptions are made about the distribution of the proportion of focus shifts. Thus, these indices are non-parametric (Stirling 2007).

The indices that rely on the calculation of proportions of focus shifts (Shannon H and SID) share a number of properties outlined by James and Taeuber (1985) and discussed by Reardon and Firebaugh (2002). The most relevant characteristics to consider in the context of multitasking measurement are Equivalence and Size Invariance. Each is explained below:

1. *Equivalence*: If a work session of a specific time length is divided into segments, whereby each one has the same proportion of focus shifts as the original session, diversity does not change. Similarly, if several consecutive session segments of smaller duration are combined into a single session, with the same proportion of focus shifts across tasks, diversity does not change.

2. *Size invariance*: If the number of focus shifts in each task is multiplied by a constant, the diversity of the resulting session is the same as the diversity of the original session. The principle of size invariance indicates that the indices are insensitive to sample sizes. For example, two sessions with different number of focus shifts will yield the same results as long as the distribution of proportions across tasks is equivalent for both sessions.

These considerations provide the basis for recommending the adoption of one index over the other. Researchers concerned with identification and measurement of dominant tasks should use order 2 measures such as SID. The higher the order, the more the index emphasizes the most common or dominant task. In contrast, order 1 measure (Shannon's H) gives equal weight to all tasks regardless of their proportions and it is maximized when there is complete evenness. Alternatively, when the frequency of population elements (proportion of focus shifts) is not important, richness is an appropriate choice.

Since indices based on proportions of focus shifts give more information than richness alone, researchers should consider using more than one measure. For example, to complement Shannon's H, and to overcome the limitation of size invariance, researchers could simultaneously consider the richness of the session as well as the number of focus shifts. For a single measure, Shannon H is the best choice due to its fairness. It counts all the tasks based on their proportion of focus shifts without favoring neither the most common, nor the rarest. This index has also been proven to be robust when extended to more complex measurement contexts such as multi-group segregation (Reardon and Firebaugh 2002).

Empirical validation

To demonstrate the validity of Shannon's H when applied to the measurement of multitasking, we used the *known-groups* technique (MacKenzie et al. 2011). This approach consists of testing whether the means of the focal measure are significantly different across groups with recognized differences in the construct of interest. In order to apply this technique, a sample of self-reported diaries of computer usage collected for the study reported in Benbunan-Fich et al. (2011a) was used. The diaries were received from an assignment embedded in an undergraduate introductory course in Information Systems taught at a large urban college located in the Northeast of the U.S.

Diaries are a method of data collection used for obtaining information about episodic or recurrent behaviors, and are particularly suitable to collect information about how time is used. Data obtained from diary-reporting is more reliable than questionnaires because measuring time use results in complex surveys. Data collection instruments that impose a high burden on participants increase the risk of biases and non-responses (Kenyon 2008). In addition, since the diary method is self-administered, it is free from interviewer and/or researcher effects. However, the insertion of diary reporting into normal activities may cause interference. Despite their potential intrusiveness, diaries have been found to provide reliable information (Juster and Stafford 1991).

Depending on the format and reporting requirements, diaries vary in their degree of structure. This research used semi-structured diaries collected for another study and stored in a database. Participants of the original study were required to keep a diary or log of a session of computer usage. For this purpose, they used a special form where they indicated with time stamps, the tasks, technology applications, and reasons for shifting to another task or application during a session. These diaries were structured to the extent that participants were given a specific form to record their computer-use behaviors. However, participants were free to explain their activities in their own words. Neither the time units nor the activities to be recorded were fixed. Hence, the diaries were semi-structured.

To minimize the problem of retrospective recall, participants filled out their diary by indicating tasks, applications, times and reasons for changing tasks, as their work session took place. This ensured a more accurate reporting of time-use activities. Because of the burden imposed by the data collection method, the requirements of diary reporting were limited to one to two hours of computer use. Limiting the time frame for diary keeping was necessary because producing concurrent documentation of longer sessions increases the risk of underreporting activities.

The database had 160 diaries with durations between 29 and 139 minutes and a median session length of 63 minutes. The majority of the diaries were concentrated in the 50 to 70 minute range. Therefore, for this study, diaries with duration of 1 hour plus/minus 10 minutes were selected. This selection procedure yielded 107 diaries with the required duration. In the original study, each diary was given a 0-1 classification by a pair of coders to indicate whether the participant was multitasking during the session. From the original coding, the only variable retained for this validation procedure was the dichotomous multitasking indicator.

Results

A different pair of coders (not familiar with the original study) was convened to calculate the number of tasks and focus shifts for each of the selected participants' diaries. Inter-coder reliability was 0.93 for number of tasks (richness). The coders resolved disagreements by discussing discrepancies and reaching consensus on a final task determination. Eleven diaries in which participants were performing a single task (single task diary sessions) were discarded. The final sample included 96 (107 minus 11) diaries. After reaching agreement on the tasks, the coders determined the number of focus shifts in each diary (Inter-Coder reliability of focus shifts is 0.86). A similar adjudication procedure was used to resolve coding differences in focus shifts. Using the agreed upon focus shift distribution for each diary, Shannon's H in its original and normalized form (H*) was calculated for each diary. Table 1 shows the descriptive characteristics of the final sample.

Table 1. Descriptive Statistics of Continuous Variables

	N	Mean	Std. Dev.	Min	Max
Session Duration	96	61.38	4.40	50.00	70.00
Number of Tasks	96	4.57	1.53	2.00	8.00
Total Focus Shifts	96	6.85	5.22	0.00	27.00
H	96	1.19	0.50	0.00	2.01
H*	96	0.79	0.27	0.00	1.00

Session duration of the diaries in the sample ranges from 50 to 70 minutes, with an average of 61 minutes. As a result of removing diaries with only one task, the number of tasks reported in the diaries of the sample varies from 2 to 8. The average number of tasks is between 4 and 5 tasks with a standard deviation of 1.53. The number of focus shifts goes from 0 to 27, with an average of almost 7. As explained in the discussion of prototypes, zero focus shifts occur when a person performs tasks in sequence, completing the first before starting the next. H ranges from 0 to 2 with a mean of 1.19 and a standard deviation of 0.50, and the standardized index (H*) ranges from 0 to 1 with an average of 0.79 and a standard deviation of 0.27.

Based on the existing dichotomous multitasking indicator in the database, about 57% (55 diaries) received one, indicating multitasking activity, and 43% (41 diaries) received a 0 because they exhibited very little to no-multitasking during the reported session. The 0-1 indicator was used to defined two groups (multitasking vs. non-multitasking). Table 2 shows the descriptive statistics of session duration, number of tasks, and focus shifts in each group. The two groups have noticeable differences in the number of focus shifts (3.39 in the non-multitasking and 9.44 in the multitasking group) but not in the average number of tasks, which is approximately the same for both (about four tasks). While the average of focus shifts is different between the groups, it is not known whether these shifts are distributed differently across tasks.

Table 2. Descriptive Statistics By Groups

	Non-Multitasking Group				
	N	Mean	Std. Dev.	Min	Max
Session Duration	41	60.78	3.76	51.00	70.00
Number of Tasks	41	4.22	1.47	2.00	7.00
Total Focus Shifts	41	3.39	2.17	0.00	8.00
	Multitasking Group				
	N	Mean	Std. Dev.	Min	Max
Session Duration	55	61.82	4.81	50.00	70.00
Number of Tasks	55	4.82	1.54	2.00	8.00
Total Focus Shifts	55	9.44	5.34	2.00	27.00

As discussed, Shannon's H provides a joint assessment of both focus shift distribution and number of tasks per session, and thus gives a measure of the extent to which multitasking behavior occurs. A statistical comparison between these two groups was computed to detect the presence of significant differences in H between these two groups. Due to the size of the sample, a non-parametric test was used. A t-test comparison between these two groups for H and H* produces equivalent results. Table 3 shows the means of H and H* in each group, descriptive statistics and the results of the inter-group comparison.

Table 3. Two Group Comparison

	H				
	N	**Mean**	**Std. Dev.**	**Min**	**Max**
Non-Multitasking Group	41	0.94	0.78	0.00	1.79
Multitasking Group	55	1.37	0.32	0.64	2.01
Kruskal-Wallis Test χ^2= 10.87**					
	H*				
	N	**Mean**	**Std. Dev.**	**Min**	**Max**
Non-Multitasking Group	41	0.62	0.34	0.00	1.00
Multitasking Group	55	0.91	0.07	0.73	1.00
Kruskal-Wallis Test χ^2= 24.73***					

Significance levels: ** p<.001; *** p<.0001

In the non-multitasking group, there were a handful of diaries with about the same number of tasks and focus shifts, such that most tasks only had a single focus shift. This pattern indicates that each task was attended only once during the session, abandoned to work on another and never resumed. The computation of H in these cases yields a high value (more than 1.6 for sessions with 6 or 7 tasks) and explains why the maximum of the non-multitasking group is 1.79. Moreover, when the index is standardized, H* reaches levels closer to 1, even though the corresponding diary is not reflective of a multitasking situation. One remedy to prevent the distortion of the index due to these particular situations is to correct the focus shifts quantification ex-post, by considering that the tasks are concluded if they are not returned to during the session. In so doing, the pattern will resemble the sequential prototype and the index will be zero or very low.

In spite of these few cases where the index is high, the average H for the non-multitasking group is 0.94, which is significantly lower than the average H for the multitasking group (1.37). This difference is significant at p<.001. For H*, the average values are 0.62 and 0.91 for the non-multitasking and multitasking group, respectively, and this difference is significant at p<.0001. Thus, according to the known-groups technique, the index successfully discriminates between multitasking and non-multitasking situations.

Discussion

This study integrates Activity Theory and the concept of focus shifts analysis to develop metrics for multitasking activity. To this end, multitasking is defined as shifts in the focus of attention when the flow of work is interrupted and tasks are switched during a session. The notion of focus shifts and the count of shifts to each task during a period of time (or session) can be used as a frame of reference. The tasks that receive attention determine the classification of focus shifts, and focus shift analysis provides the foundation for measuring multitasking.

The starting point for the calculation of any measure based on focus shift analysis is to report the total number of tasks and the total number of focus shifts in a session. Neither one represents an actual index of diversity, but each one provides a valuable description of the session. Moreover, these two parameters can be combined in a ratio and used as a basic description of the session and a preliminary indicator of multitasking activity. For example, the patterns of figures 1, 2 and 3, would be represented as 3:0; 3:6 and 3:4 respectively. Likewise, the case shown in Figure 4 would be represented as 5:15.

An advantage of this preliminary information presented in the form of a ratio (tasks to focus shifts) is its simplicity to identify multitasking situations without the need to perform additional calculations. If the number of focus shifts is either zero, very small or lower than the number of tasks, then there was little multitasking activity in a session. In particular, if the number of focus shifts is approximately equal to the number of tasks (e.g. 3:4), it means that the session is partially sequential or has embedded tasks, and deserves closer examination. Conversely, if the number of focus shifts is greater than the number of tasks (e.g. 3:6 or 5:15), a more sophisticated index is needed to calculate the extent of multitasking activity.

Although a ratio is useful as a preliminary multitasking indicator, particularly to compare sessions with the same number of tasks but different number of focus shifts, it gives incomplete information. A ratio does not provide an accurate representation of the pattern of dispersion of focus shifts during a session (i.e. how the number of focus shifts is distributed across tasks). In this case, it is necessary to systematically quantify the diversity of focus shifts. To this end, this study discussed two possibilities – Shannon's index (H) and Simpson's index of diversity (SID). These measures take into account richness (number of tasks) as well as the relative distribution of focus shifts across tasks. Shannon's index, which is considered the more robust of the two, was empirically validated with the known-groups technique (MacKenzie et al. 2011) with data collected from self-reported diaries. The results show that the index captures well the extent to which people multitask, and effectively separates multitasking from non-multitasking situations.

Limitations

Although Shannon's H tends to be more robust, it is not free from limitations. The index is sensitive to focus shifts distribution but insensitive to the actual number of focus shifts involved. Therefore, it may not perform well when two sessions with different numbers of focus shifts are compared. In addition, the index is maximized in case of complete evenness (equal distribution of focus shifts across tasks) and this may not be always desirable.

The results of the empirical validation suggest other limitations. In the multitasking group, the normalized version of the index (H*) produces values that are close to each other despite differences in tasks and focus shifts. As shown in Table 3, there are a few cases where H and H* have high values for the non-multitasking group because the pattern of focus shifts is mostly composed of ones. Once identified, these cases could be adjusted before the index is computed. In addition, for the multitasking group, the standard deviation of H* is relatively small (0.07), in spite of wider variations in task counts and total focus shifts. This observation suggests that one should use caution when comparing sessions with different number of tasks and focus shifts with the standardized version of the index.

Recommendations

Despite its limitations, diversity indices are valuable to expand the methods whereby multitasking behavior is understood and measured. Our overall recommendation is that researchers think about the objectives of their study and consider the use of these indices. From those examined here, Shannon's H is particularly suitable when evenness is the conceptual dimension of interest and sensitivity to sample size (actual number of focus shifts) is not required.

Future Research Directions

This work can be extended in multiple directions. In the area of metrics development, the proposed index H can be compared with other potential indicators of multitasking, such as those in Benbunan-Fich et al. (2011a). In the area of empirical validation, future research should examine in more detail whether the discriminant ability and the statistical comparability of these indices are appropriate, when applied to different samples, or to data collected from alternative sources such as surveys.

Conclusion

Using Activity Theory as the theoretical foundation and the notion of focus shifts analysis, this study proposes a set of indicators for multitasking. The measures range from a simple ratio to a more complex index that captures the diversity of focus shifts. The suitability of each indicator depends upon the objectives of the research and the data sources and methods accessible to the researcher. In this vein, the contribution of this study is to bring a new conceptual and methodological foundation to investigate multitasking behavior from the perspective of focus shift analysis. At the conceptual level, the expansion of the concept of focus shifts offers a solid foundation to describe and investigate multitasking behavior. At the methodological level, the proposed diversity indices, and Shannon's H in particular, allow one to measure the extent to which multitasking occurs. This novel perspective has the potential to generate a new stream of empirical studies that document multitasking patterns in various settings.

Acknowledgments

This paper is the continuation of my paper titled "An Entropy Index for Multitasking Behavior," presented in the Breakthrough Ideas track of ICIS 2011, in Shanghai. This extension describes the theoretical underpinnings of the measurement approach via focus shift analysis, compares different entropy indices, and reports the results of their empirical validation. These papers have benefited from the insightful feedback of anonymous reviewers, associate editors, track chairs, and conference attendees, who have generously shared their ideas and suggestions to improve my work. I am sincerely grateful to all of them. I am also thankful to Priya Pant and Laura Del Rio for their capable research assistance and to Melissa Fich for her valuable help in the editing of this paper.

References

Adams, W. (2006) "Wireless laptops in the classroom," *Communications of the ACM* (49:9), pp. 25-27

Baczkowski, A.J., Joanes, D.N. and Shamia, G.M. 1997. "Properties of a Generalized Diversity Index," *Journal of Theoretical Biology* (188), pp. 207-213.

Bedny, G. Z. & Karwowski, W. 2007. *A Systemic-Structural Theory of Activity: Applications to Human Performance and Work Design*. Boca Raton, FL: CRC/Taylor & Francis.

Bedny, G. Z., Seglin, M. H. and Meister, D. 2000. "Activity Theory: History, research and application," *Theoretical Issues in Ergonomics Science* (1:2), pp. 168-206.

Benbunan-Fich, R., Adler, R. and Mavlanova, T. 2011a. "Measuring Multitasking Behavior with Activity-Based Metrics," *ACM Transactions on Human-Computer Interaction* (18:2), June, Article #7.

Benbunan-Fich, R. 2011b. "An Entropy Index for Multitasking Behavior," *Proceedings of the International Conference on Information Systems (ICIS)*, Paper 2, December, Shanghai, China.

Benbunan-Fich, R. and Truman, G. 2009. "Multitasking with Laptops during Meetings," *Communications of the ACM* (52:2), pp. 139-141.

Bluedorn, A. C., Kaufman, C. F. & Lane, P. M. 1992. "How Many Things Do You Like to Do at Once? An Introduction to Monochronic and Polychronic Time," *Academy of Management Executive* (6:4), pp. 17-26.

Bødker, S. 1996. "Applying Activity Theory to Video Analysis: How to make sense of video data in Human-Computer Interaction." In Nardi, B. A. (ed.), *Context and Consciousness*. Cambridge, MA: MIT press, pp. 147-174.

Desrochers, R.E. and Anand, M. 2004. "From Traditional Diversity Indices to Taxonomic Diversity Indices," *International Journal of Ecology and Environmental Sciences* (30), pp. 85-92.

James, D. R., and Taeuber, K.E. 1985. "Measures of Segregation," *Sociological Methodology* (14), pp. 1–32.

Juster, F.T and Stafford, F.P. 1991. "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement," *Journal of Economic Literature* (29), pp. 471-522.

Kaptelinin, V. & Nardi, B. A. 2006. *Acting with Technology: Activity Theory and Interaction Design*. Cambridge, MA: MIT Press.

Kenyon, S. 2008. "Internet use and time use: The importance of multitasking," *Time & Society* (17:2-3), pp. 283-318.

Kuutti, K. 1991. “Activity theory and its applications to information systems research and development.” In H.-E. Nissen, ed., *Information Systems Research*. Amsterdam: Elsevier Science Publishers, pp. 529-549.

MacKenzie, S. B., Podsakoff, P. M. and Podsakoff, N. P. 2011. “Construct measurement and validation procedures in MIS and behavioral research: Integrating new and Existing Techniques,” *MIS Quarterly* (35:2), pp. 293-334.

McWilliams, G. 2005. “The laptop backlash,” *The Wall Street Journal*. Oct. 14, B1.

Pielou, E.C. 1975. *Ecological Diversity*. London, UK: Wiley.

Reardon, S. F. and Firebaugh, G. 2002. “Measures of MultiGroup Segregation,” *Sociological Methodology* (32:1), pp. 33-67.

Rényi, A. 1960. “On measures of entropy and information,” *Proceedings of the Fourth Berkeley Simposium on Mathematical statistics and Probability*, pp. 547-561.

Shannon, C.E. 1948. “A mathematical theory of communication,” *Bell System Technical Journal* (27), pp. 379-423.

Simpson, E.H. (1949) Measurement of diversity. *Nature*, 163: 688.

Stirling, A. 2007. “A general framework for analysing diversity in science, technology and society,” *Journal of the Royal Society Interface* (4:15), pp. 707-719.

Waller, M.J. 2007. “Preferences, Behaviors, and Strategies in Multiple-Task Performance,” in Dansereau, and Yammarino (ed.) *Multi-Level Issues in Organizations and Time (Research in Multi-Level Issues, Volume 6)*, Emerald Group Publishing, pp. 239-247.

Wallis, C. 2006. “The multitasking generation,” *Time Magazine* (167:13), Mar. 27, pp. 48–55.

Watson, J. M. and Strayer, D. L. 2010. “Supertaskers: Profiles in extraordinary multitasking ability,” *Psychonomic Bulletin & Review* (17:4), pp. 479-485.