Estimating Behavior Frequencies:

Do Vague and Enumerated Estimation Strategies Yield Similar Results?

James S. Cole

Ali Korkmaz

Indiana University – Bloomington

Paper presented at the annual meeting
of the Association for Institutional Research
May 23, 2011
Toronto, Ontario, Canada

Introduction

Estimating behavior frequency is difficult, but important in survey research. Surveys that collect data regarding behavior estimates are found in many fields including, but not limited to, those that conduct consumer research, health studies, sexual behavior, drug use, political polls, and many types of education studies. These studies typically use either vague behavioral quantifiers as the response set ("sometimes", "often", etc), or enumerated response sets where the respondent needs to select or tally the target behavior (e.g., "This past semester, how many times did you meet with your advisor?"), or a combination both types. There is evidence that vague quantifiers have some advantages over enumerated (tallied) responses in that they are less cognitively taxing and therefore more easily answered (Lenzner, Kaczmirek, & Lenzner, 2010). As a result, they also tend to be associated with a higher survey completion rate. However, they also are more difficult to interpret (Wright, Gaskell, & O'Muircheartaigh, 1994). For instance, given the difficulty of interpreting what is meant by "often" how do we know that "often" is more than "sometimes" for any given behavior? This study investigates the relationship between vague and enumerated quantifiers for a set of educationally-relevant behaviors.

Literature Review

Uses of Vague and Enumerated Quantifiers

Surveys of student behavior commonly collect data regarding estimations of behavior frequency in two ways. One way is to ask the respondent to estimate behavior using vague quantifiers (e.g., "often"). For instance the *National Survey of Student Engagement* (NSSE)

published by the Center for Postsecondary Research at Indiana University asks respondents, "During the current academic year, about how often have you . . . worked with other students on projects during class." Vague response categories include "never," "sometimes," "often," and "very often." Another example is the survey *Beginning Postsecondary Students Longitudinal Study* conducted by the National Center for Education Statistics (US Department of Education) which includes items such as "how often respondent wrote essay answers as part of exams during the 2003-2004 academic year" with response categories "never," "sometimes," and "often." The other way to estimate behavior is the use of enumerated quantifiers where the respondent either selects or enters a numeric value that best represents the estimated frequency of the target behavior. For instance, on the *Beginning College Survey of Student Engagement* published by the Center for Postsecondary Research at Indiana University, respondents are asked "During your last year of high school, about how many hours did you spend in a typical 7-day week doing each of the following?" and then provided with a list of activities and numeric response categories (0, 1-5, 6-10, 11-15, etc). Similarly, the *University of California Undergraduate Experience Survey* published by the University of California, asks respondents "How much time do you spend in a typical week (7 days) on the following activities . . . attending classes, discussion sections, or labs". The ordered response categories include 0 hours, 1-5 hours, 6-10 hours, etc. Overall, a review of commonly used higher education surveys provides many examples of both types of response sets. Vague and enumerated response sets can also be found on many other higher education surveys.

Though collecting survey data regarding estimates of behavior are common practice, they are not without controversy (e.g., Porter, Rumann, & Pontius, in press). Researchers have

noted serious limitations for interpreting data collected using both types of response sets (Schaeffer & Presser, 2003; Wanke, 2002). For instance, researchers have noted the difficulty when interpreting what is meant by each vague responses category (what is really meant by "often"?), whereas it is well-established that enumerated responses are prone to many sources of error in the process of recall (Bradburn & Miles, 1979; Brown, 1995).

Referencing Versus Enumerating When Estimating Behavior Frequency

The process of estimating behaviors using vague quantifiers is distinct from the process used to estimate behavior by trying to tally or count occurrences. Prior research has shown that questions with vague response sets invoke a process of social or behavioral comparison to estimate the behavior frequency (Pace & Friedlander, 1982; Schaeffer, 1991; Wanke, 2002). Though the process goes by different names, for instance Sudman, Bradburn, and Schwarz (1996) refers to respondents comparing their behavior to "anchors" for comparison, where as Wright, Gaskell, and O'Muircheartaigh (1994) refers to the self information hypothesis where a person compares his or her behavior to an internal "benchmark" and the group norm hypothesis where "comparative information used in the response process is derived from the behaviors of the social groups to which the person belongs"(p. 481-482), there is agreement that the process for responding to a vague quantifier involves the process of comparison. In other words, how to estimate how often a respondent participates in an activity always occurs within the context of "often" compared to what or whom? For instance, in a study by Wanke (2002), students were asked how often they go to the movies. One group of students was told the study examines leisure activities of city residents (general population); whereas the second

group was told the study investigates the leisure habits of their fellow university students. Students were asked how often they go to the movies using a vague quantifier and then later an enumerated response. Results showed no significant difference in the enumerated response of how often they went to the movies; however there were significant differences in the vague responses between the two reference groups. The study found that “often” was relative to the frequency of the behavior exhibited by the comparison group, whereas the enumerated response was an estimated tally of behaviors not influenced by the comparison group. Other research has shown that respondents also use similar activities, expectations, or other reference points as their comparative anchors when selecting a vague quantifier to estimate their behavior (see Schwarz & Oyserman, 2001).

The other process for estimating behavior is by tallying or counting the occurrences for a set time frame. For instance, “How many times have you met with your academic advisor this current academic year?” This type of question will require the respondent to tally the number of occurrences within a given time frame. This enumeration process involves one or two steps depending on the time frame, behavior frequency, and schedule (fixed or variable) (Brown, 2002; Schaeffer & Presser, 2003). The first step is simply enumeration. Low frequency behaviors regardless of the time frame or schedule, generally involves a straightforward tally of occurrences and are easily recalled. For example, most students can generally count how many times they failed a test during current academic year. For most students this is a very rare occurrence and, to the extent it happened, they could probably tell you about it with a good deal of accuracy. However as the frequency of the behavior increases, the time frame and schedule become critical variables in the recall process. For instance, tallying how often you

asked questions in class over the course of an academic year may be quite difficult for most students. When the respondent is asked to tally such a question, at least two steps are needed to provide a response. The typical strategy involves shortening the time frame to a point where it can be tallied ("I asked questions in class 8 times this past week") then extrapolate to the longer time frame ("Each semester is 15 weeks, we are in week 10 of the second semester for a total of 25 semesters, therefore I asked questions about 200 times since the beginning of the school year"). Brown (2002) refers to this two-step process as enumeration and extrapolation, and as noted by Brown is "more common as presentation frequency increases" (p. 41). However, given the cognitively taxing task of providing enumerated estimates of frequent behaviors over long periods of time, there is evidence that respondents use a strategy referred to as "satisficing" (Blair & Burton, 1987). Schaeffer and Presser describe satisficing as the process of "conserving time and energy and yet producing an answer that seems good enough for the purposes at hand" (p. 68). The result of satisficing is the clumping of numerical estimates around common multiples such as 5 or 10 (Krosnick, Narayan, & Smith, 1996; Huttenlocher, Hedges, & Bradburn, 1990).

Though enumerated responses have the appearance of accuracy, it is important to note that enumerated responses are not necessarily accurate. As pointed out by Brown (2002) and Schaeffer (1991), the accuracy of enumerated responses is dependent on many factors such as satisficing, and the accuracy of stored and recalled information. From cognitive psychology, we know that there are many factors associated with effective encoding and/or effective retrieval of information (Ashcraft & Radvansky, 2009). Thus forgetting or error in memory can be caused by decay of memory traces over time, interference from other related memory traces,

repression (especially for negative events), construction error where the information recalled is inadvertently constructed erroneously which leads to inaccurate recall, situational and environmental cues, and temporary failure to retrieve (Ormrod, 2008). Another important factor regarding enumerated responses is that the accuracy of recalled events presupposes that the individual is making a good faith effort to accurately recall information. For instance, it is well established that men report two to four times more opposite-sex partners as women. It is suspected that men often intentionally misrepresent the number of sex partners resulting in social desirability bias (Brown & Sinclair, 1999).

Given the many difficulties involved with enumerating past behaviors, especially for frequent behaviors over long periods of time, some researchers have suggested that relative, vague judgments of behavior are preferred (Bradburn & Danis, 1984). Sudman, Bradburn and Schwarz (1996) go so far as to say, "Since behavioral frequency reports are error-prone anyway, why bother asking respondents for reports that suggest more precision than they can provide?" (p. 226). However, Sudman, Bradburn and Schwarz (1996) go on to say, "vague frequency expressions carry their own load of problems. . . different respondents use the same term to mean difference objective frequencies of the same behavior" (p. 226). Thus, it is not surprising that researchers often find group differences for meanings assigned to vague quantifiers (e.g., Wanke, 2002).

Group Differences in Estimating Behavior

Due to the comparative process involved when a respondent interprets vague response sets, a concern with some researchers is that students interpret vague quantifiers differently

based on their background, social groups, or past experiences. As pointed out by Schaeffer (1991), “when the meaning of vague quantifiers differ by group, relative and absolute responses may lead to different conclusions” (p. 397). This norming effect becomes particularly problematic when examining groups differences using responses to vague quantifiers (Wright et al., 1994). The literature to date generally provides evidence to suggest that there may be some important group differences in the way that vague response categories are interpreted. However, not all studies have found these differences. Pace and Friedlander (1982) found that the greatest variation was between activities, and not between groups of students. Pace and Friedlander found that “occasionally” reflected a different number of events between writing and use of the library. However, they also did find some group differences, particularly by grade level (mostly B+ or better, mostly B’s, mostly B- or lower) and by science and non-science majors. They also found that there were significant differences between selective, private institutions and larger comprehensive, doctoral institutions. However, given the small number of institutions, they did not see any need to develop separate norms by institution type. In another important study, Wright et al (1994) found a significant interaction between the age of the respondent and gender regarding their perceptions of how often they watch television. They also found that interpretation of vague quantifiers differed significantly between socio-economic classifications. They conclude that, “people differ systematically in how they interpret a response scale composed of vague quantifiers” (p. 493). Other studies have found similar results. For example, Schaeffer (1991) found significant differences in the meaning of the vague response categories based on race, education, and age, but no differences between males and females. However, Brown and Sinclair did report gender differences when estimating the

number of sexual partners (1999). Differences by age were noted by Borgers, Hox, & Sikkel (2003). In another study, Bradburn and Miles (1979) found differences in interpretation of vague quantifiers depending on the respondent's affect (positive or negative affect or pleasant or unpleasant emotions, respectively). A study by Nelson Laird, Korkmaz, & Chen (2009) found significant differences in the interpretation of vague quantifiers and enumerated behaviors for gender and major, however these differences were small. Except for the Pace and Friedlander study in 1982, there has been very few (if any) published studies that investigated behavioral estimations by groups of students regarding educationally relevant behaviors. This study set out to provide additional evidence regarding group differences regarding estimation of important educationally-relevant behaviors.

Overall, all the studies described above lead researchers to wonder if "often" is only meaningful when the context is known, and/or the respondents' background and experiences are known. For instance, when a respondent answer a question such as "How often did you join a study group in the past 12 months", the researcher does not know with whom the respondent is comparing his or her response ("often" compared to her friends?), or with what she is comparing their response ("often" compared to studying by herself?), or some other reference point that may be related with the respondents gender, ethnicity, or institution type, or some other characteristic. This study extends this line of research by investigating the following questions:

1. Is there a relationship between the frequency of the behavior and the time frame selected to estimate the frequency of the behavior?

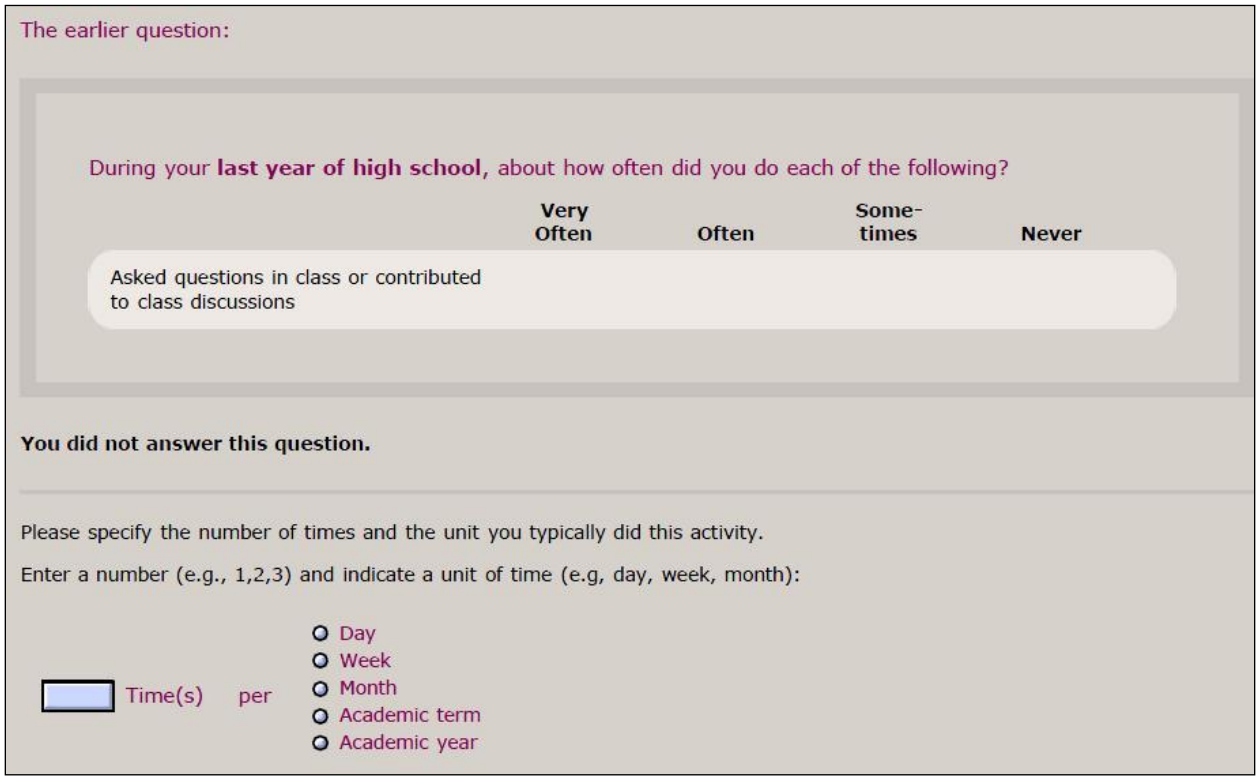
2. For each behavior is there an increasing enumerated estimation with each increasing level of vague quantifier?
3. Are there group differences between vague and enumerated estimates of the same behaviors based on gender, type of high school (public/private), or completion of calculus (an indicator of academic achievement level)?

Method

Data Source

Data for this study comes from the 2010 administration of Beginning College Survey of Student Engagement (BCSSE). Many of the items in BCSSE use vague quantifiers. The project routinely uses both paper and Web survey modes. For this study, 6 items from the core survey were repeated at the end of the Web version of the survey for students at a subgroup of institutions. Students were reminded of their original response to the item and were then asked to quantify their response by indicating how many times they did the activity per day, week, month, academic term, or year. Students filled in the amount and selected the time frame (see Figure 1). The content of the 6 items focused on six high school behaviors (see Appendix A for complete listing of items).

Figure 1. Example questionnaire page soliciting absolute frequency associated with vague quantifier.



Sample

Respondents included 30,964 first-year entering students from 81 institutions representing a range of types. These include 12 Doctoral, 39 Master's, 24 Baccalaureate, and 6 "other" types of institutions. About 57% (46) were private institutions; and 10 % were minority serving institutions. Institutional size representation was 37% were small (less than 2,500 students), 25% medium (2,500-4,999), 22% large (5,000-9,999) and 16% very large. Student characteristics mirrored those of all students participating in BCSSE in 2010. About 41% of the respondents were male. Of the responding students, 67% were white, 8% were African American, 9% were Latino, and 8% were Asian or Asian American. About 37% of the

respondents were first-generation students and 17% of the students attended private schools. About 36% of students passed Calculus in high school.

Data Analysis

Student responses to the enumerated items were recoded so that all responses were on a per week basis. Students could report frequency based on five options (per day, week, month, academic term, and academic year). Week was taken as the baseline and responses with other time frames were adjusted by appropriate multipliers (day = 5, month = .25, academic term = .111, and academic year = .02778). There are 180 days in a typical high school academic year, 4 terms of 9 weeks.

Analyses were run to answer each of the three specific research questions. For the first research question, we used the enumerated items to determine the relationship between high and low frequency activities and time frames chose. We rank order the frequency of the activities and create a frequency distribution for each item to determine patterns.

In order to test the distinct representativeness of the vague response options by enumerated response options, we examined the distribution of specific quantities by each vague quantifier and survey item. For each item, we run separate ANOVA test to check the distinctiveness of vague quantifiers and used omega squared. Omega squared is an estimate of the dependent variance accounted for by the independent variable for a fixed effects model.

To examine whether the meaning of response options varied by student characteristics, we used t-test and Cohen's effect size to detect mean differences in both formats. For each of the six items, we calculated the mean differences in vague and enumerated formats. We checked if one method favors one group of students or not by checking the difference and

direction of the difference. Student characteristics included gender, type of high school (public/private), and completion of calculus (an indicator of academic achievement level). We summarized the findings for each student characteristics and overall.

Results

1. *Is there a relationship between the frequency of the behavior and the time frame selected to estimate the frequency of the behavior?*

Yes. Table 1 below indicates the rank order for each behavior with the most frequent behavior being “ask questions” and the least frequent being “class presentations” (The rank orders were calculated by standardizing the time frame where the enumerated estimates for each behavior were converted to a weekly estimate and then put in order from most frequent to least frequent). Results show that the more frequent the, the shorter the time frame the respondent uses when estimating behavior frequency. Approximately 68% of the respondents used “day” as the time frame when estimating the number of times they asked questions in class (an overall high frequency activity), whereas a little more than 3% of the students chose “day” when estimating the number of presentations (an overall low frequency activity). Likewise, 50% of the respondents used “month” as the time frame when estimated the least frequent behavior (making a class presentation).

Table 1. Rank order of frequency of behaviors by time frame used to estimate behaviors

Frequency Rank order	Behavior	Time Frame Used to Estimate Behavior Frequency				
		Day	Week	Month	Term	Year
1	Ask questions	67.9%	27.9%	2.6%	0.5%	1.1%
2	Discussed ideas w/ others	21.9%	44.1%	20.9%	5.3%	7.9%
3	Discuss grades/assign w/teachers	9.9%	40.1%	31.0%	12.3%	6.8%
4	Discussed ideas w/ teacher	11.2%	32.8%	27.0%	10.1%	18.9%
5	Worked w/classmates outside class	5.3%	28.3%	35.3%	15.9%	15.2%
6	Class presentations	3.3%	26.6%	49.9%	12.0%	8.1%

2. *For each behavior is there an increasing enumerated estimation with each increasing level of vague quantifier?*

Yes, main effects were all significant with medium to large effect sizes (Omega squared). Table 2 shows that each increasing level of vague quantifier is associated with higher levels of enumerated behavioral estimations. However, the large standard deviations indicate a problem with the distribution of the data (see Table 2). Additional analysis revealed significant interactions with time frame and enumerated behavioral estimations for all items. In addition, upon further analysis it was discovered the multivariate assumptions could not be met. Levene's Test for equality of error variance was significant ($p<.001$) for all all DV's. Therefore, a Games-Howell post hoc test was conducted. The Games-Howell test is used when variances are unequal and when there are unequal group sizes. The Games-Howell test does not assume a equal variances. The Games-Howell test reduces the chance of Type I error due to unequal variances, especially due to unequal group sizes (Maxwell & Delaney, 2000). To test whether

there is an increasing enumerated estimation for each level of vague quantifier, a one-way ANOVA was run for each level of the independent variable. Since the independent variable has four levels, a Bonferroni adjustment was calculated to reduce the chance of a Type I error. The adjustment indicated that with the significance level at .01 for each level, the family-wise error rate was kept below .05.

As seen in Table 3 for “Ask questions”, there were significant differences between each level of vague quantifier and the associated mean enumerated response for those students that chose the most common times of Day and Week (68% chose day and 28% chose week). This result indicates that each vague level of response for this question is associated with a significantly higher level of enumerated estimation. The most commonly chosen time frames for the other items was week and month. Using these time frames we see similar results are found for all other items. Each level of vague response is associated with significantly higher level of enumerated response. The one exception was for class presentation where the difference between sometimes and often was not significant.

Table 2. ANOVA results for vague quantifiers items by enumerated responses

		Weekly Mean	N	SD	df	F	sig	Omega sq
Ask	Never	2.32	329	5.3	3	2142.4	.001	.185
	Sometimes	7.36	5913	8.5				
	Often	15.31	10338	12.4				
	Very Often	22.86	11709	14.9				
Class Pres	Never	0.80	518	0.3	3	1997.8	.001	.173
	Sometimes	0.49	11459	0.5				
	Often	0.98	11930	0.8				
	Very Often	1.30	4725	0.9				
Faculty outside class	Never	0.16	7042	0.7	3	4222.2	.001	.309
	Sometimes	1.09	13456	1.5				
	Often	2.67	5570	2.3				
	Very Often	4.01	2313	2.9				
Discuss grades	Never	0.17	1259	0.7	3	2272.5	.001	.192
	Sometimes	0.89	12181	1.3				
	Often	2.14	10528	2.1				
	Very Often	3.19	4667	2.6				
Students Outside Class	Never	0.11	3115	0.5	3	3073.3	.001	.243
	Sometimes	0.71	14954	0.9				
	Often	1.66	7943	1.4				
	Very Often	2.26	2670	1.6				
Others Outside Class	Never	0.30	2259	1.3	3	4016.4	.001	.300
	Sometimes	1.63	12198	2.1				
	Often	4.07	9420	3.3				
	Very Often	6.55	4166	4.2				

Table 3. Games-Howell post-hoc test results.

			DAY	WEEK	MONTH	TERM	YEAR
Ask questions	Never	a	8.9	2.1	0.5	$0.4^{b,c,d}$	0.0
	Sometimes	b	14.9	3.5	1.0	$0.6^{a,c,d}$	$0.3^{c,d}$
	Often	c	20.4	5.4	2.2^{d}	$1.0^{a,b,d}$	$0.4^{b,d}$
	Very Often	d	26.0	7.6	3.7^{c}	$2.5^{a,b,c}$	$0.7^{b,c}$
Class presen	Never	a	2.1	0.9	0.2	0.1	0.0
	Sometimes	b	$10.4^{c,d}$	1.8^{d}	0.5	0.3	0.1
	Often	c	$12.5^{b,d}$	1.9^{c}	0.7	0.5^{d}	0.2
	Very Often	d	$12.3^{b,c}$	2.3	0.8	0.6^{c}	0.3
Faculty grades	Never	a	0.8	1.2	0.4	0.2	0.0
	Sometimes	b	$6.5^{c,d}$	1.9	0.5	0.3	0.1
	Often	c	$6.9^{b,d}$	2.5	0.8	0.4^{d}	0.2^{d}
	Very Often	d	$7.0^{b,c}$	3.1	1.0	0.5^{c}	0.3^{c}
Students outside	Never	a	0.2	1.3	0.3	0.1	0.0
	Sometimes	b	$4.7^{c,d}$	2.0	0.5	0.3	0.1
	Often	c	$5.0^{b,d}$	2.4	0.8	0.4	0.2^{d}
	Very Often	d	$5.0^{b,c}$	2.9	1.0	0.6	0.3^{c}
Discuss ideas w/fac	Never	a	0.5	1.5	0.3	0.1	0.0
	Sometimes	b	6.5	2.0	0.5	0.3	0.1^{c}
	Often	c	7.0^{d}	2.7	0.9	0.5^{d}	$0.2^{b,d}$
	Very Often	d	7.4^{c}	3.4	1.1	0.7^{c}	0.3^{c}
Discuss ideas w/others	Never	a	1.3	1.7	0.4	0.1	0.0
	Sometimes	b	7.2	2.3	0.6	0.3^{d}	0.1
	Often	c	8.1	3.3	1.1	0.6^{d}	0.3^{d}
	Very Often	d	9.1	4.6	1.6	$0.7^{b,c}$	0.4^{c}

Note - Superscript indicates non-significant differences between indicated level(s). No superscript indicates significant difference with other level(s).

3. *With each increasing level of vague quantifier are there groups differences for enumerated estimates of the same behaviors based on gender, type of high school (public/private), or completion of calculus (an indicator of academic achievement level)?*

For each set of items, we calculated mean differences and effect size for vague quantifier and enumerated formats for gender. As shown in Table 4, two out of six items showed the same pattern in both formats (discuss w/ faculty and others outside class) meaning if male students were shown higher in one format, the other format displayed a similar trend for these two items. However, the other four items showed different patterns. For example, "asking questions in class", "discuss grades with faculty", and "discuss with students outside class" favor female students in vague format but favors male students in enumerated format. Though the differences are significant in these three items, the effect sizes are small.

We ran the same analyses for high school type. As shown in Table 5, four out of six items showed the same pattern in both formats of the survey responses. Effect size difference was comparatively higher in only one of these items (ask questions in class). "Discuss grades with faculty" item favored public schools in vague format but in enumerated format favored private schools. However, the effect size difference was very small in both formats. "Class presentation" item favored vague quantifier format but not significant in enumerated format. Moreover, the effect size difference was in the same direction and very close to each other.

We ran the same analyses for high school calculus takers. As shown in Table 6, again four out of six items showed the same pattern in both formats of the survey responses. Again, the same item (class presentation) showed significant difference in vague quantifier but not in enumerated format and effect size difference very similar in both formats. "Discuss with faculty

outside class" reversed the direction of difference for groups but not significant in enumerated format.

Table 4. Group differences between vague and enumerated estimates of the same behaviors based on gender.

		Gender			
		Male	*Female*	*sig*	*d*
Ask questions	Vague	3.09	3.16	.001	-.088
	Enumerated	16.81	16.47	.043	.024
Class presentations	Vague	2.65	2.75	.001	-.128
	Enumerated	0.82	0.81	.253	.014
Discuss grades w/faculty	Vague	2.66	2.69	.001	-.037
	Enumerated	1.78	1.63	.001	.073
Discuss w/studs outside class	Vague	2.29	2.37	.001	-.097
	Enumerated	1.07	1.04	.026	.027
Discuss w/faculty outside class	Vague	2.13	2.12	.032	.015
	Enumerated	1.50	1.34	.001	.076
Discuss w/others outside class	Vague	2.45	2.62	.001	-.194
	Enumerated	2.87	3.21	.001	-.097

Table 5. Group differences between vague and enumerated estimates of the same behaviors based on completion of type of high school.

		High School Type			
		Public	*Private*	*sig*	*d*
Ask questions	Vague	3.13	3.19	.001	-.080
	Enumerated	16.16	19.46	.001	-.235
Class presentations	Vague	2.71	2.69	.026	.022
	Enumerated	0.82	0.81	.121	.024
Discuss grades w/faculty	Vague	2.69	2.66	.004	.028
	Enumerated	1.67	1.80	.001	-.064
Discuss w/studs outside class	Vague	2.31	2.41	.001	-.126
	Enumerated	1.02	1.20	.001	-.141
Discuss w/faculty outside class	Vague	2.10	2.25	.001	-.175
	Enumerated	1.35	1.67	.001	-.156
Discuss w/others outside class	Vague	2.53	2.63	.001	-.124
	Enumerated	3.01	3.35	.001	-.100

Table 6. Group differences between vague and enumerated estimates of the same behaviors based on completion of calculus.

		Calculus			
		No	*Yes*	*sig*	*d*
Ask questions	Vague	3.10	3.20	.001	-.116
	Enumerated	15.86	17.97	.001	-.149
Class presentations	Vague	2.70	2.68	.037	.017
	Enumerated	0.82	0.81	.118	.019
Discuss grades w/faculty	Vague	2.71	2.62	.001	.110
	Enumerated	1.74	1.59	.001	.074
Discuss w/studs outside class	Vague	2.28	2.40	.001	-.153
	Enumerated	1.01	1.12	.001	-.082
Discuss w/faculty outside class	Vague	2.11	2.14	.001	-.043
	Enumerated	1.41	1.37	.110	.020
Discuss w/others outside class	Vague	2.51	2.62	.001	-.124
	Enumerated	2.97	3.26	.001	-.083

Discussion

This study set out to investigate the relationship between two methods for estimating educationally related behaviors. Overall this study found that the more frequent the behavior, the shorter the time frame the respondent uses when estimating the behavior. This finding is consistent with previous literature (Brown, 2002). The implication for researchers is that time frame is an important factor when asking respondents to enumerate estimates of past behaviors.

Another important finding is that this study showed that vague quantifiers for many educationally relevant behaviors are associated with increasing enumerated responses for the

same behavior. This provides additional reliability evidence for the use of vague quantifiers. In this case, two behavioral estimates are providing consistent estimations of the behavior. Still unanswered is the accuracy of these estimates, but these data do provide additional evidence regarding the reliability of the estimates.

Another equally important finding is that there were minimal group differences regarding these estimates. Though there were many significant differences by gender, high school type, and calculus completion, the effect sizes for all differences were very small. The significant difference findings were likely due to the large sample size in this study. A single institution performing the same study with a much small sampler may find very few significant differences. The implication of these results is that for these types of behaviors there are likely no meaningful differences between these groups of students and how they interpret vague and enumerated estimates of behavior.

References

Ashcraft, M. H., & Radvansky, G. A. (2009). *Cognition*. Upper Saddle River, NJ: Prentice Hall.

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research, 14*, 280-288.

Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of labeled response options and vague quantifiers. *International Journal of Public Opinion Research, 15*, 83-94.

Bradburn, N. M., & Catalina D. (1984). Potential Contributions of Cognitive Research to Survey Questionnaire Design. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds). *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, pp. 101-29. Washington, DC: National Academy Press.

Bradburn, N. M., & Miles, C. (1979). Vague quantifiers. *Public Opinion Quarterly, 43*, 92-101.

Brown, N. R. (1995). Estimation strategies and the judgment of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1539-1553.

Brown, N. R. (2002). Encoding, representing, and estimating event frequencies: A multiple strategy perspective. In E. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition,* pp. 37-53. Oxford: Oxford University Press.

Brown, N. R., & Sinclair, R. C. (1999). Estimating number of lifetime sexual partners: Men and women do it differently. *Journal of Sex Research, 36,* 292-297.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 196-213.

Kinzie, J., Gonyea, R., Kuh, G. D., Umbach, P., Blaich, C., & Korkmaz, A. (November, 2007). *The relationship between gender and student engagement in college*. Paper presented at the Annual Conference of the Association for the Study of Higher Education, Louisville, KY.

Krosnick, J. A., Narayan, S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation, 70*, 29-44.

Lenzner, T., Kaczmirek, L., & Lenzner, A. (2010). Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied Cognitive Psychology, 24*, 1003-1020.

Maxwell, S. E., & Delaney, H. D. (2000). *Designing experiments and analyzing data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum.

Nelson Laird, T. F., Korkmaz, A., & Chen, P. D. (2009). *How often is "often" revisited: The meaning and linearity of vague quantifiers used on the National Survey of Student Engagement*. Paper presented at the Annual Conference of the American Educational Research Association, San Diego, CA.

Pace, C. R., & Friedlander, J. (1982). The meaning of response categories: How often is "Occasionally," "Often," and "Very often"? *Research in Higher Education, 17*, 267-281.

Ormrod, J. E. (2008). *Human Learning*. Upper Saddle River, NJ: Pearson.

Porter, S. R., Rumann, C., & Pontius, J. (in press). The validity of student engagement survey questions: Can we accurately measure academic challenge? *New Directions for Institutional Research.*

Schaeffer, N. C. (1991). Hardly ever or constantly: Group comparisons using vague quantifiers. *Public Opinion Quarterly, 55*, 395-423.

Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology, 29*, 65-88.

Schwarz, N., & Oyersman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation, 22*, 127-160.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996), *Thinking about answers.* San Francisco, Jossey-Bass.

Wanke, M. (2002). Conversational norms and the interpretation of vague quantifiers. *Applied Cognitive Psychology, 16*, 301-307.

Wright, D. B., Gaskell, G. D., & O'Muircheartaigh, C. A. (1994). How much is 'Quite a bit'? Mapping between numerical values and vague quantifiers. *Applied Cognitive Psychology, 8*, 479-496.

Appendix A.

BCSSE 2010 items

Six of the items are from the section that asks respondents about their high school experiences. The items below are numbered to correspond with the survey numbering. Students responded to these questions in a vague format in the core survey, and then with an enumerated format as additional items added at the end of the survey (see Figure 1 as example).

BCSSE 2010 core survey items

9. During your last year of high school, about how often did you do each of the following?

a. Asked questions in class

b. Made a class presentation

d. Discussed grades or assignments with a teacher

f. Worked with classmates outside of class to prepare class assignments

i. Discussed ideas from your readings or classes with teachers outside of class

j. Discussed ideas from your readings or classes with others outside of class (students, family members, etc)