Abstract

This thesis examines the function of adverbs of quantification through a comparison of questions found in two public health surveys, the data collected from each survey, and the subsequent allocations of healthcare resources and focus on specific areas in the medical field. After examining the types of questions present in the Adverse Childhood Experience (ACE) survey and the US Household Food Security Scale Module (HFSSM), I conclude that each of the surveys utilizes adverbs of quantification differently in their questions. The ACE survey’s main objective is to establish the psychological impacts of childhood trauma, so while the possible answers may be linguistically vague, the survey collects important data to inform mental health professionals about how to treat patients. The HFSSM establishes more specific possible answers to create a scale of food security that can be nationally applied and assist the US government in allocating resources to fight food insecurity. I argue that survey authors would benefit from focused, basic linguistics training or a linguistic advisor to help write questions that avoid presuppositions and implicit restrictions, and help researchers collect the most helpful data possible.

1. Introduction

This thesis first examines previous studies of adverbs of quantification in order to establish a framework for an investigation into public health surveys. Without prior knowledge of how adverbs of quantification operate, healthcare professionals in charge of writing surveys will lack an understanding of the mechanisms involved in comprehending their questions, which can have
serious consequences on the kinds of data they want to collect and what conclusions they draw from that data.

This thesis studies adverbs of frequency, which provide information about the quantity of their constituents, rather than making the constituent itself more specific. I examine adverbs of quantification that do not assign a fixed quantity to their constituents, such as often, sometimes, rarely, and ever, which I will interpret as “at least once.”

1.1. An issue with syntax

Because questions in health surveys can be written in different styles, it is important to consider potential problems we face when considering the syntax of an interrogative that contains an adverb of quantification. Henriette de Swart (1993) says that when studying determiners, arguments are usually assumed to be fixed by the syntactic structure of a sentence: the determiner of a subject NP takes the denotation of the subject noun as its first argument, and the set of individuals denoted by the VP as its second argument. When studying adverbs of quantification, syntax is less helpful in fixing the quantifier’s arguments, since adverbial position is not as strict as that of determiners. de Swart lists (1) with different readings (2-3):

(1) Anne always knits Norwegian sweaters.
(2) All of the sweaters that Anne knits are Norwegian sweaters.
(3) Anne knits a Norwegian sweater every time she knits.

With more than one way to map restriction to scope, it is vital to establish the ways in which adverbs of quantification choose their arguments in order to avoid ambiguity. The first option de Swart mentions is that the restriction on the adverb of quantification comes from the sentence itself. (2) and (3) demonstrate this – in (2), the fact that Anne knits sweaters implies she knits Norwegian sweaters, and in (3), where the fact that Anne knits something suggests she knits
Norwegian sweaters. Other sentences utilize a temporal adverbial and the second argument of the main clause to determine the first argument of the quantifier, as seen in (4-6):

(4) In the evening, Anne always knits Norwegian sweaters.

(5) When he is ill, Paul is mostly in a bad mood.

(6) Before he goes to bed, Paul sometimes takes a shower.

Here, the adverb of quantification quantifies over the temporal clause and considers the number of situations which extend to the type of event described by the main clause. (4) is true only if all contextually relevant evenings overlap with circumstances in which Anne knits Norwegian sweaters. de Swart’s third option is an implicit restriction on the quantifier with context being the restriction source, (7-9) demonstrate:

(7) Anne mostly goes to Spain.

(8) Anne always refuses.

(9) Anne never brings a present.

At first glance, and out of any context, these sentences seem odd, but in the right setting they can be used without any issues. (7) fits into a discussion about people taking a vacation: in most of the cases where Anne takes a vacation (the restricted domain), she goes to Spain. (8) includes the presupposition that it is significant that Anne always refuses, and thus her utterance of refusal is meaningful (Potts 2014); an appropriate context could be that every time Anne is asked to wash dishes, she refuses. A sensible context for (9) could be that it is expected of housewarming party guests to bring a gift, but every time Anne is invited to a housewarming party, she doesn’t bring a gift.

Restriction can also be implicit, although it appears that most surveys do well avoiding this. I did not find any evidence of questions with implicit restriction in a health survey, but it is not
difficult to imagine a scenario in which it can arise. A health survey could ask subjects to rate their agreement with the statement “I felt pain every time.” Right away, it sounds like a strange question – the overt restrictor for the quantifier every is time, or each event of the subject feeling pain, but the question isn’t really asking about literally every single time the subject ever felt pain. Instead, it is asking about a specific circumstance in which the subject felt pain – the domain is being restricted, and the restriction is coming from something that isn’t in the sentence. It could be coming from the context in which the survey is administered, or from the manner in which the subject is considering the question; in any case, it is not made explicitly clear what the context of the question is, which can cause the data collected to be less informative towards the original inquiries of the health professional giving the survey.

This section is especially important to the survey part of this thesis, since these kinds of healthcare questions are often context-dependent and thus more difficult to control for when writing and scoring a survey. As I will explain later, this syntactic variation of restricting adverbs of quantification rarely appears on health surveys, and if it does, there is usually a great deal of context provided for the question.

1.2. Relational and non-relational readings

Continuing with her theme of context, de Swart next discusses relational and non-relational readings. Adverbs of quantification can have different semantic and pragmatic meanings depending on the situation to which they are referring. de Swart illustrates this with the sentences in (10) and (11):

(10) Paul often goes on a holiday.

(11) Paul often goes to the movies.
and (11) could easily have different numbers of occurrences for their events, even if they happen “often” in both contexts. (10) might be true if Paul goes on a holiday more than 3 times a year, while (11) might be true if Paul goes to the movies 3 times a month. This case of vagueness is easy to solve, since we know that in life, people go to the movies more often than they go on vacation. I will go deeper into this discussion later, but this is obviously a critical element of healthcare patients’ answers on surveys. A certain experience might have occurred “often” for two different people, but without knowing if the circumstances in which the event occurred are similar or different for each person, the exact frequency of the experience is impossible to estimate, and thus is not very helpful when using that data to make what is assumed to be an informed decision.

1.3. Distinguishing the object of quantification

In his work to figure out exactly what object was getting quantified, David Lewis (1975) listed six groups of quantificational adverbs. While this thesis will not consider all of these adverbs, they are listed for breadth of coverage in (12-17):

(12) always, invariably, universally, without exception

(13) sometimes, occasionally [once]

(14) never

(15) usually, mostly, generally, almost always, with few exceptions, [normally]

(16) often, frequently, commonly

(17) seldom, infrequently, rarely, almost never

Lewis’ first guess was that adverbs of quantification quantify over times. This seems reasonable when looking at the first adverbs in each list in (12-17); for example, the sentence

Always $\Phi$ - adverb and the original sentence $\Phi$ - is true if and only if the original sentence is true
at all times. However, Lewis points out that the times quantified over can be periods instead of
moments, as demonstrated in (18):

(18) The fog usually lifts before noon here.

The sentence in (18) that is modified by *usually* is true for most days, not most moments in
time, because days are the relevant context here. I will discuss context and its function briefly in
the next section.

Lewis also writes that the range of quantification can be restricted by the information in a
sentence, using (19) as an example:

(19) Caesar seldom awoke before dawn.

This modified sentence is not concerned with all the times Caesar was alive; it considers all
the times Caesar awoke, and says that of those times, few of them were before dawn.

Lewis then considered the idea that adverbs of quantification quantify over events. After
considering (19), it might seem like the adverbs being studied quantify over events, with
acceptable restriction. However, an event can be a continuous state instead of a single
occurrence, as Lewis shows in (20):

(20) A man who owns a donkey always beats it now and then.

This means that beatings occur over the course of the entire ownership of the donkey, but the
act of ownership isn’t usually considered an event by itself. (20) also works to defeat guess 1: if
*always* quantified over time, we would have dueling modifiers that suggest both that the beatings
happen without pause (*always*) and that the beatings happen infrequently (*now and then*).

Lewis concludes that adverbs of quantification are quantifiers over cases: what is *sometimes*,
*never*, *often*, or *seldom* true is true in *some*, *no*, *many*, or *few* cases respectively. He also
introduces the idea of adverbs of quantification as *unselective* quantifiers that bind all variables
within their scope. I would expect to see these unselective quantifiers appear in questions on public health surveys as part of a strategy to make comprehension easier.

After establishing a basic understanding of adverbs of quantification, I now shift to examining how medical surveys are created.

2. **Questionnaire and study design**

Researchers take many steps in designing a clear and comprehensive health survey, carrying out data collection, performing analyses, and drawing conclusions about healthcare trends. There are many steps in this process, and many different entities that can have input. Before studying actual surveys, it is critical to understand the methods behind their creation.

It is important to recognize that the questions in public health surveys are not always completely new. Questions are sometimes taken from previously published surveys and articles, with phrasings that are often quite similar, if not identical. Definitions of medical conditions or incidents can also be taken from or influenced by past works. In a study using the Adverse Childhood Experience survey (1995), Felitti et al. (1998) utilized seven different sources to provide definitions of childhood trauma, including five surveys, one article, and one mental health screening diagnostic. It is difficult to create entirely new questions for health surveys without inspiration, and this “borrowing” makes it easier to continue with a “status quo” of certain types of questions and keep their phrasing over years and even decades of use. This can lead to resistance to change, even when evidence is presented that other ways of asking questions may tell researchers more important information than what was previously being gathered. Debates about survey wording are usually settled with pilot testing – a sort of practice run to make sure everything researchers want to happen is happening, and catching any surprises.
Health surveys often must go through rounds of protocol editing before getting approved for use on the public. There are several levels of jurisdiction that survey creators might have to consider. A local health network that will be the actual proctors of the survey must be consulted, as it should have the best feel for the patients it serves. Regional medical schools may want to provide input and have access to study results as part of their own research. Government agencies like the National Institute of Health or the Center for Disease Control and Prevention may want to modify surveys to collect data that can be compared with previous studies. In short, there are many different influences on the creation and execution of a public health survey. My examination of questions from two different surveys in the next section explores the challenges of using public health data to establish national policies while simultaneously providing necessary attention to local populations directly involved in the surveys. I will also discuss how asking different kinds of questions yields different data that may be useful in ways that are not purely number-driven, such as informing healthcare professionals on the psychological impacts of people’s environments.

While much of the existing literature that instructs and guides the creation of health surveys does not focus explicitly on linguistics, some work has been done in the last few years to update it. Most articles and guides to survey design focus on the usual validity and reliability standards, and they are usually concerned with the more minute details of a survey item, such as writing it as a question to answer or a statement to agree with to varying degrees, or the number of response options (Magee et al. 2013). However, I did find one instance of a linguistic analysis of a questionnaire as part of a comprehensibility assessment. The Evaluative Linguistic Framework (ELF) was developed to assess healthcare texts based on functional linguistic theory, examining how language functions as a meaning-making system, taking context into account as
well as considering how survey subjects construct meaning from text. The ELF has several metrics for determining if a survey is acceptable, grading elements like lexical density, the writer/reader relationship, and rhetorical elements as “clear/unambiguous” or “uncertain”. Clerehan et al. (2016) use it to evaluate the Health Literacy Questionnaire and find that the ELF is effective in determining the strengths and weaknesses of a questionnaire and suggesting methods of improvement. I would like to see the ELF and other similar linguistic analyses applied to widely-used surveys, including the ones I mention in the next section.

3. Questions in healthcare surveys

In this section, I examine certain questions included in the Adverse Childhood Experiences (ACE) Questionnaire, as well as in the US Household Food Security Scale Module (HFSSM).

3.1. The ACE questionnaire

The ACE Questionnaire, administered between 1995 and 1997 by the Centers for Disease Control and Prevention, asks questions about participants’ traumatic experiences during the first 18 years of their life. The questionnaire was part of a large longitudinal study that tracked participants’ health and social problems throughout their adult life. I examine some questions from the survey, drawing from the previous sections to illustrate the effect of adverbs of quantification. The question in (21) includes the adverb of quantification ever, which I have previously defined as “at least once”:

(21) Were your parents ever separated or divorced?
    Yes    No

In (22), context becomes important, as the adverb of quantification often appears:

(22) Did a parent or other adult in the household often...
    Swear at you, insult you, put you down, or humiliate you?
    or
    Act in a way that made you afraid that you might be physically hurt?
    Yes    No
This is a more complex question than the one in (21). *Often* can have different meanings depending on who is answering the question; if two different people took the ACE survey and answered “Yes,” researchers would know only that two people considered at least one of these events to happen “often” during the first 18 years of their life. The participants could easily be considering different contexts when answering this question: one could have grown up with these traumatic experiences and considered 6 incidences per month to be “often,” while the other could have been mostly removed from these experiences and judged 2 incidences per month to be “often.” It would be difficult to get exact numbers of incidences from adults trying to recall their childhood trauma, so this type of question gets asked a lot on surveys — in a way, it’s asking if the rate of incidence was significant to the participant.

The ACE survey questions get more complex with the question in (23), which contains both *often* and *ever*:

(23) Did a parent or other adult in the household *often*...
   Push, grab, slap, or throw something at you?
   or
   **Ever** hit you so hard that you had marks or were injured?
       Yes  No

While this isn’t a confounding question, it is partially context-dependent, and the question in (24) is even more complicated:

(24) Was your mother or stepmother:
   **Often** pushed, grabbed, slapped, or had something thrown at her?
   or
   **Sometimes** or *often* kicked, bitten, hit with a fist, or hit with something hard?
   or
   **Ever** repeatedly hit over at least a few minutes or threatened with a gun or knife?
       Yes  No

This is interesting for a couple of reasons. We see the new adverb of quantification *sometimes*, and we see that it has been differentiated from *often*. That means survey participants
have to consider two different possible contexts for the same childhood events: were they kicked, bitten, etc. “some” of the time, or “much/most” of the time? As was the case in (22), different participants will have different frames of reference, and thus can give the same “Yes” answer while having vastly different childhood experiences. Conversely, participants can also have very similar childhood experiences and give very different answers. I discuss the implications of this in section 4.

3.1.1. A note on presupposition

(21) offers a chance to examine the dangers of presuppositions in health survey questions. A model for the inferences of the question is below, using “R” to represent the researchers who created the survey and “S” for the subject completing the survey:

(25) a. R/S: Were your parents ever separated or divorced?
   b. (i) CONVEYS: S had knowledge of S’s parents.
      (ii) CONVEYS: R believes S had knowledge of S’s parents.
      (iii) CONVEYS: S’s parents were married.
      (iv) CONVEYS: R believes S’s parents were married.

(25a) can only be asked if it is established for both the researchers and the subject that the subject had parents that were married and that the subject had knowledge of their parents’ marital status, otherwise the answers provided will not reflect the subjects’ actual childhood living situations. That presupposition was not very risky for the original population of the ACE survey: Per CDC data, the survey was administered to volunteers who were members of a San Diego health maintenance organization (so all subjects had jobs and healthcare); 74.8% of the subjects were white and 75.2% of them had attended college. In short, it was not a stretch for the researchers to presuppose that their subjects knew their parents and that their parents were married, since that knowledge was assumed to be part of the common ground for most of the subjects. Of course, that is not always the case: people can grow up with other caregivers, and
they can also grow up with parents that aren’t legally married. (25) would not be an effective question to ask these kinds of people, since it is asking about a particular circumstance that never existed in their lives, and answering it at all would be too much to accommodate (Stalnaker 2002). (21) also does not address the issue of biological parents; in some cases it may be that a subject grew up with foster parents after being adopted. In short, (21) may not lead to accurate answers, which will obviously be detrimental to the goals of the survey.

Other questions in the ACE survey eliminate some presuppositions. (23) asks about parents or an “other adult,” and (24) asks about a mother “or stepmother,” signaling that the question designers considered potential caregiver circumstances that may not have included biological parents. As I mentioned in section 2, it is important for people who write surveys to have at least a rudimentary understanding of semantics and pragmatics; a linguistic advisor or even basic linguistics training can help avoid potentially confusing wording that gives rise to confounding presuppositions.

3.1.2. Data from ACE questionnaire

Several United States health organizations have run surveys similar to the ACE questionnaire in recent years and collected data from hundreds of thousands of people in order to make informed decisions about policies, divide supplies and financial aid, and elevate public awareness of national health issues. The Substance Abuse and Mental Health Services Administration, a branch of the U.S. Department of Health and Human Services, analyzed the ACE data and made a number of influential conclusions:

*ACEs are common.* 28% of participants reported physical abuse, and 21% reported sexual abuse. Many participants reported divorce or parental separation, or having a family member with a substance abuse or mental disorder.
ACEs often occur together. Almost 40% of the participant pool reported at least 2 ACEs, and 12.5% reported 4 or more. After this finding was published, many subsequent studies focused more on the cumulative effects ACEs have on long term health, rather than focusing on each experience’s individual outcome.

ACEs have a dose-response relationship with many adult health problems. As the number and frequency of ACEs increases, participants’ health problems in adulthood multiplied and had earlier onset times. These health problems are unfortunately extremely diverse.

- Substance-related issues: more ACEs predict earlier initiation of alcohol use (Rothman et al. 2008), continued tobacco use during adulthood (Ford et al. 2011), and increased likelihood of illicit drug use (Dube et al. 2003).

- Behavior problems: ACEs increased the likelihood of lifetime depressive episodes (Ege et al. 2015), sleep disturbances in adults (Kajeepeta et al. 2015), less healthy infants at birth (Smith et al. 2016), more long-term health issues such as diabetes or heart attacks (Monnat & Chandler 2015), and poor dental health (Bright et al. 2015).

3.2. The Household Food Security Scale Module (HFSSM)

The HFSSM was originally implemented in 1995 by the US Department of Agriculture as part of the US government’s initiative to develop a standard for measuring food insecurity at local, state, and national levels. Updated modules appear every few years, as it plays a primary role in the well-being of the US population (Bickel et al. 2000). As in section 3.1., I examine some questions from the survey, drawing from the previous sections to illustrate the effect of adverbs of quantification. HFSSM questions like (26) are very similar to the ACE question (21):
(26) In the last 12 months, did you \textbf{ever} cut the size of your meals or skip meals because there wasn't enough money for food?

Yes
No
DK

This is simply asking the participant to identify if this marker of food insecurity was present in their lives at any time in the past 12 months. The difference between the ACE survey and the HFSSM is that the HFSSM uses follow-up questions to further clarify participants’ answers.

(27) How often did this happen—almost every month, some months but not every month, or in only 1 or 2 months?

Almost every month
Some months but not every month
Only 1 or 2 months
DK

The HFSSM questions appear to provide more informative answers than the ACE survey questions, establishing a scale on which to determine food insecurity. The sequence of (26) and (27) is an example of a set of survey questions in which the issue about dependency on context raised in section 1.1 is eliminated. The writers of the HFSSM provide a monthly domain that is much more specific than the ACE survey, which is interested in the much broader “first 18 years” of the subjects’ lives. This strategy does make for longer surveys, which can be a strain on both the subjects taking surveys and the researchers analyzing them, but more specific information creates a more informative picture of subject populations that can be quickly translated into funding or resource allocation.

3.2.1. Data from the HFSSM

The HFSSM provides enough data to extrapolate to the entirety of the United States. In 2017, 88.2% (112.3 million) of US households were food-secure, 7.3% of US households had low food security, and 4.5% of US households had very low food security. At the state level, the
prevalence of food insecurity was statistically significantly higher than the national average in 11 states, as shown in Figure 1:

![Prevalence of food insecurity, average 2015-17](image)

**Figure 1.** Prevalence of food insecurity in the 50 states. Data from 2015 to 2017 was combined to provide more accurate representation of state trends.

Because the supplemental programs allocate federal resources to states based on need, it is vital to have current and accurate data for food insecurity by state. Thus, the HFSSM aims to collect the most specific information possible from its participants.

4. **Final conclusions and future directions**

The HFSSM and ACE survey have questions that share a similar structure and use of adverbs of quantification, but the researchers who administer them have different goals for each one. The ACE survey questions are attempting to determine how people think and feel about traumatic childhood experiences they may have endured, not necessarily what the exact experiences were or exactly how often they occurred. For the implications of the survey – a greater emphasis on therapy and recognizing/preventing childhood trauma – questions with adverbs of quantification that yield potentially vague answers still accomplish the goal of the research. In contrast, the goal of the HFSSM is to establish a countrywide database of nuanced information that informs
national agricultural and food aid policies and resource allocations. Thus, the module must have questions with adverbs of quantification that eliminate vagueness as much as possible.

While these surveys clearly collect useful data and are, for the most part, well-written, there are still some linguistic elements that could be improved. Researchers should do their best to either gain some understanding of the semantics and pragmatics of questionnaires, and potentially also consult with a medical linguistics expert when possible, to maximize the effectiveness of their surveys. Materials like the ELF exist and are applicable for some surveys; a more widely-applicable framework is a likely next step.

Sometimes, surveys can be poorly written and yield vague data, but they can still be helpful for researchers to find more specific data trends to investigate or learn what to improve for future surveys. The Inspiration Scale (Thrash and Elliot 2003) is a psychological survey that measures inspiration, asking subjects to rate “how often” they experience an event on a scale from 1 (never, or not at all) to 7 (very often, or very strongly/deeply). The survey includes the following four questions:

(28) I experience inspiration.
    How often does this happen? 1 2 3 4 5 6 7
    How deeply or strongly (in general)? 1 2 3 4 5 6 7

(29) Something I encounter or experience inspires me.
    How often does this happen? 1 2 3 4 5 6 7
    How deeply or strongly (in general)? 1 2 3 4 5 6 7

(30) I am inspired to do something.
    How often does this happen? 1 2 3 4 5 6 7
    How deeply or strongly (in general)? 1 2 3 4 5 6 7

(31) I feel inspired.
    How often does this happen? 1 2 3 4 5 6 7
    How deeply or strongly (in general)? 1 2 3 4 5 6 7
At first glance, this may look like a survey that isn’t very helpful for a direct approach at helping a patient feel more inspired. Questions abound: How is (28) different from (29) if they both ask about individual events of experiencing inspiration? What’s the difference between being inspired to do something, in (30), and feeling inspired, in (31)? This survey certainly wouldn’t accomplish the same task that the ACE and HFSSM surveys aim to accomplish; the answers — the data that is collected — simply doesn’t tell much about the subject beyond how they feel inspiration. There isn’t even a domain restriction written in the survey, so how is it helpful to the health professionals who administer it? The answer is that while data can only tell you a person’s answer to a question, they can suggest other pathways to investigate in an attempt to uncover other factors in a person’s life that are closely related to, but not examined by, the answers to a survey. A patient in a psychologist’s office might provide a “5” answer to (29), but a “1” answer for (31); the seemingly strange dichotomy of the two answers provides a talking point in the therapy session that may uncover something about the patient’s understanding of inspiration, which the psychologist can use to tailor their treatment. A person who answers the ACE survey with evidence of multiple adverse experiences, but still presents well in adulthood, may have many reasons for their resilience, but the survey can’t know about them because it did not ask about them. Survey data doesn’t give researchers all of the answers they might need to solve a problem — it only gives one set of answers. Thus, it is important for researchers to be deliberate in their study design, as they can adjust their focus to give data that can either prompt a quick response or suggest a deeper look into a population of study subjects.

In this thesis, I have performed a comparative analysis of two surveys that differ in question format, as well as the populations studied. While this has clearly been informative, another type of study may provide even more insight into the effect of survey question structure.
Urke et al. (2014) cite growing economic uncertainty and the importance of food security as a
determiner of health as reasons for investigating whether a shorter, less comprehensive
questionnaire could be used adequately to identify food insecurity. Their test subjects were
members of an Inuit tribe in Artic Canada, so their results can’t be directly compared with the
US population, but they are still interesting: overall, single-question inquiries increased
sensitivity to food insecurity (the researchers had more success determining who was food-
insecure) but decreased specificity (the researchers had a more difficult time determining who
was not food-insecure). If developed further, this kind of rapid assessment technique could be
useful for healthcare and social service providers hoping to quickly identify food-insecure
households, and it could potentially be integrated into larger populations without much strain on
data collectors. An ideal test of a rapid assessment technique might involve giving both a full-
length and shortened HFSSM to two groups from the same population and comparing the data to
find significant or potentially problematic discrepancies.

Acknowledgements
I would like to thank Theodore Fernald and Shizhe Huang of the Swarthmore and Haverford
Linguistics Departments, respectively, for their advising during the writing process, as well as
Sierra Bienz and Lydia Brunk for their feedback. I am also grateful for Donna Jo Napoli, who
was the original inspiration for my linguistic career.
References


Potts, C. 2014. Presupposition and implicature. Handbook of Contemporary Semantics 2. 168-


