A behaviorist correspondence theory of truth

By: Emil Alexander

C-thesis in Theoretical Philosophy
Uppsala University
Department of Philosophy
Supervisor: Tobias Alexius

Introduction

For many decades there has been an ongoing feud between the fields of behaviorism and cognitive science. This feud is not about specific scientific findings, it is about deep philosophical convictions, and about what terms and methods it makes sense to use when studying psychology. In the late 1950's, behaviorism was declared dead when it was convincingly argued that behaviorism could not explain the nature of language, a centerpiece of human psychology. But since then behaviorism has slowly risen from its grave, as a new behaviorist theory of language emerged. The new behaviorist theory of language is called Relational Frame Theory (RFT), and it is part of a new behaviorist paradigm called Contextual Behavioral Science (CBS). This paradigm also includes a behaviorist psychotherapy called Acceptance and Commitment Therapy (ACT), which in the last decade has become popular across the world. Thus, the feud has once again become active, and the question about which philosophical principles are most suitable for the science of psychology is yet again something that needs an answer. But things have changed since the mid-1900's when the discussion was last active. The philosophy of CBS is not exactly like that of earlier versions of behaviorism, having developed into a more explicit and coherent set of philosophical principles, summarized under the name functional contextualism. Old arguments against behaviorism do not apply to the same degree. So it is time for a new look at this debate, taking into consideration what functional contextualism and RFT has to offer.

According to Contextual behavioral science, cognitive science generally entails a commitment to the correspondence theory of truth, the idea that something is true if it corresponds with reality, or a worldly fact. CBS, on the other hand, makes an explicit commitment to a pragmatic theory of truth, which focuses on the consequences (i.e. usefulness) of a statement or theory, instead of its correspondence with reality. Because of the supposed centrality of these theories of truth for the divide between cognitive science and behaviorism, I will focus on what exactly this divide is about, and whether there is any way that the differences can be reconciled. I will argue that the divide isn't as big as it may seem when we take a closer look at the philosophical principles and empirical theories of CBS, and that it may in fact be possible to bridge this divide by formulating a version of the correspondence theory that is compatible with CBS.

In part 1 I will present a quick sketch of behaviorism as contrasted with cognitive science, and the connection between behaviorism and the pragmatic theory of truth, as well as the connection between cognitive science and the correspondence theory of truth. In part 2 I will give a more detailed description of the philosophy and science of Contextual behavioral science, including the tools for understanding language in CBS terms. In part 3 I will present a more detailed description of the correspondence theory of truth, giving an overview of the different versions of this theory that have been proposed throughout the history of philosophy. In part 4 I will make a careful evaluation of the CBS objections towards the correspondence theory of truth, and arrive at a version of the correspondence theory that can be expressed in CBS terms. I will conclude that this version is compatible with the underlying philosophy of CBS, even though the CBS pragmatic theory of truth claims otherwise.

Part 1: The feud between behaviorism and cognitive science, and how it relates to the question of truth

1.1 The history and basics of behaviorism and cognitive science

The feud between the fields of behaviorism and cognitive science has a long history. Behaviorism is a field that focuses on the study of behaviors (in a broad sense), and tries to establish fundamental principles that allows for both the prediction and control/influence of any behavior. Most famously, behaviorism concerns the principles of respondent (or Pavlovian) conditioning, and operant (or instrumental) conditioning. It studies in particular how organisms learn new behavior through association, and reinforcement and punishment, respectively. Cognitive science, on the other hand, is a field that focuses on the study of the mind (in a broad sense) and its processes. It is quite a broad field that includes the study of perception (e.g. vision, hearing), cognition (such as attention, memory, and decision-making), personality, social interactions, neuroscience and even artificial intelligence. It focuses on information; how we process and represent it. The nature of the feud between behaviorism and cognitive science concerns what terms or methods or philosophical assumptions are required for psychology to be a proper scientific endeavor.

The feud began in the early 1900's. At that time, introspective psychology was the dominant paradigm of the newly established science of psychology. Introspective psychology may be said to be an earlier version of cognitive science because it shared many of the same principles. As the name suggests introspective psychology consisted of scientists investigating their own experiences and trying to uncover structure and processes of the mind, even of the central nervous system. This paradigm failed miserably mainly because of the unreliable nature of this introspection, and the inability to replicate the findings. In response to introspective psychology, John B. Watson (1913) attempted to establish a new scientific paradigm for studying psychology that would be more properly scientific, i.e. closer to the way that the successful natural sciences (physics, chemistry and biology) were being done. This paradigm was called methodological behaviorism. It centered around the principle of respondent conditioning, discovered by Ivan Pavlov, that explained how reflexive behavior could occur in response to stimuli. Methodological behaviorism was based on the philosophical principles of empiricism, developed in particular by philosophers like John Locke and David Hume. The idea was that knowledge is the result of observations, that knowledge is only possible through sensory experience, in contrast with the rationalist idea that knowledge can come from pure reason or pure introspection that is independent of experience. With this empiricist view came an attitude of suspicion against unobservable or hypothetical entities or events.

Thus, a fundamental idea of behaviorism is that psychology should only concern itself with observable events. Hypothetical or inferred entities such as beliefs, mental states, or representations were not considered scientific. One such example is the attempt to explain a behavior in terms of a person's personality. How do we know what a person's personality is? We cannot directly observe it. We can only infer the personality by observing actual behaviors. So a person who frequently acts kindly is said to be a "kind" person, to have a "kind" personality. But if we later try to explain why the person acts kindly in a certain instance by referring to that person's personality, then we have not added anything of value. All we would be saying is that the person is now acting kindly because it has acted kindly in similar situations before. Thus, the previous observable behaviors are enough. Similarly, talk about beliefs is not useful either, according to behaviorism. If a kid is told that a bear eats blueberries, and is later asked what bears eat, they will likely reply "Blueberries!". When trying to explain this, it doesn't add anything to say that the kid had a belief that bears eat blueberries. It is sufficient to refer to the fact that the kid had previously been convincingly instructed that bears eat blueberries.

Out of methodological behaviorism grew radical behaviorism, originated by B.F. Skinner (1938), which developed and focused on operant conditioning, while still retaining respondent conditioning as an important principle. Operant conditioning explained purposeful behavior, behavior that was affected by its consequences. With it, behaviorism could now explain a lot more of human and non-human psychology. With radical behaviorism the behaviorist philosophy also started shifting. Radical behaviorism was inspired by the idea of pragmatism, most notably outlined by Charles Sanders Pierce, William James and John Dewey. This was the idea that what mattered about concepts was their practical effects rather than their content. It was usefulness rather than some inherent truth that was important. This fit really well with the nature of operant conditioning.

Another change concerned that of observability. Methodological behaviorism had the constraint that all behaviors had to be *publicly* observable. That meant that no internal behaviors such as thoughts, emotions or sensations were considered scientific and should not be considered in behaviorist research. This was one of the major points of criticism against introspective psychology. With radical behaviorism, however, this view started to change, with private events becoming acceptable as scientific phenomena at least in theory. As long as one person was able to observe the phenomena (i.e. the person him-/herself), they were considered scientific. This meant that the term "behavior" expanded to include more psychological phenomena. With these philosophical and scientific developments, behaviorism flourished. Doing psychology during this time of the 40s and 50s meant doing behaviorist studies. Little cognitive science was to be seen. It seemed that behaviorism had championed. But there was one big catch: how could these principles explain human language?

In 1957, Skinner wrote a book where he made an attempt to explain how language could be understood in behaviorist terms. He used the principles of respondent and operant conditioning to explain how we learn words and how these can be used for different purposes. Language was no more mysterious than the behaviors of rats learning to press a lever for food. Despite interesting conclusions, the book did not have the desired effect on the field of psychology. A couple of years after the release, the linguist Noam Chomsky wrote a scathing review of this behaviorist attempt at explaining language (Chomsky, 1959), explaining how behaviorism failed and a cognitive paradigm was necessary in order to explain language. Among the most convincing criticisms was that of creativity of language. Every day, people are using sentences that they haven't used before, and therefore must have come up with themselves. But with respondent and operant conditioning, all behavior had to be directly trained, and thus these principles couldn't explain such creativity. Also, the behaviorist attempt could not explain how we learn grammar, since it concerned itself mostly with explaining the use of single words or phrases, but not how structure could arise among words. Chomsky reasoned that there must therefore be some cognitive explanation for these phenomena, i.e. that there must be some internal, innate computational mechanisms for learning and understanding language. The cognitive revolution started and behaviorism dwindled in popularity. Cognitive fields of psychology such as those mentioned above started to develop, and after a few decades, the roles had been reversed. Doing psychology now meant doing cognitive science.

But behaviorist research continued under the radar. In 1971, a startling discovery was made that seemed to be a key in understanding language on behaviorist grounds (Sidman, 1971). The research showed how a human via behaviorist means could learn to treat words as symbols for objects, and how this learning allowed a human to derive new behavior that it hadn't been explicitly taught. From this a new behaviorist theory of language emerged, called Relational Frame Theory (RFT), and the

phenomenon of linguistic (or verbal) behavior, was called *relational framing*. The theory co-developed with a new behaviorist philosophy, and psychotherapy, and the new behaviorist paradigm was named Contextual Behavioral Science (CBS).

Contextual behavioral science, originated by the psychologist Steven C. Hayes, grew out of radical behaviorism, retaining much of its philosophical and scientific principles, but also making its own changes in these areas. While the scientific foundation remains that of respondent and operant conditioning, CBS focuses specifically on a third behavioral principle, namely that of Relational Framing. The philosophical inconsistencies of radical behaviorism are believed to have been ironed out, and the philosophy as a whole has been explicitly outlined in a more coherent and comprehensive manner, resulting in the philosophical framework called functional contextualism. It's main point is that we have to look at the context and function of a behavior to understand it properly, but it also contains a broad set of philosophical ideas. Contextual behavioral science also includes a behaviorist psychotherapy called Acceptance and Commitment Therapy (ACT), which in the last decade has become a popular psychotherapy across the world. With the gained interest in the psychotherapy of ACT, the interest in the underlying philosophical and scientific principles has grown. The idea of behaviorism once again taking over the field of psychology is looking more like a possibility, and so the feud between behaviorism and cognitive science has been revived.

Some arguments from previous debates may still be valid, but many may be outdated because of the philosophical and empirical advancements. It is time for a new look at this debate, taking in consideration what functional contextualism and RFT has to offer. I will look at this debate from the perspective of behaviorism, or more specifically, CBS. Knowledge of the new behaviorist theories are still not well known in the field of cognitive science, and so the extant literature on the current feud is mostly written by researchers active in CBS.

Since CBS has emerged from the more general behaviorism (in particular radical behaviorism), there are some things that are true for behaviorism more generally, and some that only apply to CBS more specifically. In what follows I will use the term "behaviorism" to refer to ideas that are widely accepted within behaviorism in general, while I will use the term "CBS" or "from a CBS-perspective", when I want to be clear that it pertains specifically to CBS.

1.2 Different worldviews of psychology, according to CBS

Although Contextual behavioral science has its roots in previous behaviorist science and philosophy, the philosophical view of CBS (functional contextualism) was formulated more specifically in reference to the ideas of a philosopher named Stephen Pepper (1942). As portrayed by Hayes and colleagues (Hayes & Brownstein, 1986; Hayes, Hayes & Reese, 1988) Pepper describes four different types of worldviews, dubbed "mechanism", "formism", "organicism", and "contextualism", each based on a so-called "root metaphor" that captures the essence of each worldview. The root metaphor of mechanism is viewing the world as a machine, with discrete parts working together. The root metaphor of formism is viewing the world in terms of similarity, in the sense of similar forms recurring throughout the world. The root metaphor of organicism is viewing the world in terms of a living, or growing organic process. Lastly, the root metaphor of contextualism is viewing the world in terms of the *act in context*, meaning the behavior as understood in terms of how it relates to things outside the act itself. Each worldview is also said to have its own truth criterion, in other words, its own view on the nature of truth. In mechanism the truth criterion is correspondence, more specifically

described as some mental copy of the world being true if it corresponds with the world itself (which we will look at more closely in the next section). In formism, the truth criterion is correspondence as well, but on a simpler level of something being true if it has a form that is similar to that which it refers (for example, it is true that an object is a doughnut if it shares the form of other doughnuts). In organicism, the truth criterion is coherence, which means that when a network of facts converge and don't contain internal contradictions, then it is said to be true. Lastly, in contextualism, the truth criterion is pragmatism or usefulness, the term preferred by Hayes being that of "successful working", that is, something is said to be true if it works for the present purposes.

When Hayes and colleagues applied these worldviews to the field of psychology (Hayes & Brownstein, 1986; Hayes, Hayes & Reese, 1988) behaviorism was considered a type of contextualist worldview. But not just any type of contextualism. Hayes argued that a behaviorist should adopt a *functional* contextualist view, in line with the radical behaviorist view, in contrast with a "descriptive contextualism", meaning that the focus should be on the function that a behavior or stimulus serves, instead of what the behavior looks like. Cognitive science, on the other hand, (also called mentalism) was considered a type of mechanistic worldview, where it is considered that each part (phenomenon) can be studied in isolation, and then separately look at how these parts (phenomena) relate to each other. Thus, as functional contextualism is explained and defended, it is usually done so by being pitted against these other worldviews, and in particular against mechanism i.e. mentalism or cognitive science. From the standpoint of functional contextualism, cognitive science is therefore seen as synonymous with the worldview of mechanism, inextricably tied to its root metaphor and the correspondence theory of truth.

Although it can be of some value to ground the debate in these worldviews for the purposes of understanding differences at a fundamental philosophical level between paradigms, it can be problematic if one focuses too narrowly on these broad, generalized worldviews, since other varieties may exist beyond these four categories. Very little of the literature on functional contextualism is actually directed towards texts or claims by philosophers or scientists in the field of cognitive science. When the criticism instead is directed towards a metaphor there is a risk of the criticism being directed towards a strawman, that is, a version of cognitive science that is misrepresented and more easily attacked.

With that said, from the perspective of contextual behavioral science, one of the most central disagreements between behaviorism and cognitive science thus concerns their respective theories, or criteria, of truth. The truth criteria of the different worldviews are seen as the key feature of each worldview. Let us therefore take a closer look at the correspondence theory of truth, and the pragmatic theory of truth, respectively.

1.3 Correspondence vs pragmatic theory of truth

At least according to CBS, cognitive science entails a commitment to the correspondence theory of truth, the idea that something is true if it corresponds with reality, or a worldly fact. An example of such a theory is one that states that an idea is true if it accurately mirrors the way things really are. The way that the theory is thought to be connected with cognitive science, in the view of functional contextualism, is that cognitive science (or mechanism) concerns itself with such things as representations and mental states, which are thought of as requiring a commitment to the idea that these must correspond to things outside the mind, or some objective reality. For this reason, the term

"ontological truth" is often used as a synonym for the correspondence theory of truth in CBS literature (Hayes Barnes-Holmes & Wilson, 2012). This view may be illustrated by the following quotes:

> "[According to mechanism] the knower knows a copy of the world, not the world itself. Truth is a matter of how well the copy corresponds to the world, as evaluated by corroboration among independent knowers. Corroboration is required because the correspondence between the copy and the world cannot be observed directly." (Hayes, Hayes & Reese, 1988).

> "The truth criterion of mechanism is correspondence between the model of the world and the world as it actually is" (Levin, Twohig & Smith, 2016).

This CBS-portrayal is one possible portrayal of the correspondence theory, but other portrayals are also possible, as we will see in part 3, which may affect the effectiveness of the criticism from CBS. The correspondence theory of truth has a rich history, dating back all the way to Aristotle, and since then many different versions of this theory have been proposed. The correspondence theory of truth seems close to what we intuitively think of when we think of truth, and is also the most popular theory among philosophers (see Bourget and Chalmers 2014). But far from all philosophers agree that this is the most adequate theory of truth.

Contextual behavioral science, on the other hand, makes an explicit commitment to a pragmatic theory of truth, which focuses on the usefulness of a statement or theory, instead of its correspondence with reality. The earliest versions of a pragmatic theory of truth were proposed and defended most notably by philosophers C.S. Pierce and William James (Pierce, 1878; James, 1907). The different versions of this approach to truth are even more varied than those of the correspondence theory, but what they have in common is that they focus more on the practical matters concerning truth, such as the practical effects of claiming that something is true. Although pragmatic theories of truth aren't very popular among philosophers today, it is an integral part of CBS, and is thus gaining ground among psychologists. In part 2 I will explain in more detail the particular pragmatist stance of CBS.

Because of the supposed centrality of these theories of truth for the divide between cognitive science and behaviorism, I will focus on what exactly this divide is about, and whether there is any way that the differences can be reconciled. I will argue that the divide isn't as big as it may seem when we take a closer look at the philosophical principles and empirical theories of CBS, and that the correspondence theory of truth is actually compatible with CBS.

I will next describe in more detail the nature of Contextual behavioral science, including both its philosophy and scientific theories, and how these connect with the pragmatic theory of truth. I will then describe the correspondence theory in more detail, before contrasting the two perspectives and formulating a CBS-compatible correspondence theory of truth.

Part 2: Contextual behavioral science and the pragmatic truth criterion

This section will give an in-depth view of Contextual behavioral science and philosophy. We will begin by going through the basic terms of behaviorism (stimulus and response) and the basic principles of behaviorism (respondent conditioning and operant conditioning). This will enable us to understand the general idea of behaviorism, and what kind of phenomena and principles that functional contextualism (the philosophy of CBS) refers to. We will then explore the core terms of functional contextualism (function and context), which are needed for any CBS analysis. Then we will

be ready to properly describe the different components of functional contextualism. These components will make up the core of the objections that will later be made against the correspondence theory of truth from a CBS perspective. After this we will end with describing the new scientific theory of language and cognition of CBS: relational frame theory (RFT). This will provide us with the tools for expressing things concerning language and thinking in CBS-terms, so that when we investigate whether there is a version of the correspondence theory of truth that is compatible with CBS we can do this by attempting to express a version of the correspondence theory in such CBS terms.

2.1 The basic terms of behaviorism: Stimulus and response

The fundamental terms that are used in behaviorism are "stimulus" and "response", and in particular the "stimulus-response relation". Where response is a synonym to behavior. In behaviorism, stimuli and responses aren't considered interesting to study in their isolation or own right. Only when studying how they relate to each other in different situations do they make sense.

But what exactly is a stimulus? And what is a response? These categories are so broad that it is hard to define in any simple way. The general idea is that a stimulus is really just anything that can be observed, such as a sound, a visual object, or a sensation. A response is likewise any type of behavior, whether voluntary or involuntary, such as shouting, digesting food, thinking a thought, reacting with pain or reaching out our hand. One way of thinking about stimuli and responses are therefore that a stimulus is anything that can elicit a response, and a response is anything that can be elicited by a stimulus. One may object that these are circular definitions (as Chomsky did, 1959), but I don't think that these are understood as definitions in the proper sense. It seems it is more to illustrate the basic principle of behaviorism is that it is only in the connection between the two that they become useful.

2.2 The basic principles of behaviorism: respondent and operant conditioning

2.2.1 Respondent conditioning

Respondent conditioning (also called classical or Pavlovian conditioning), is the simplest of all behavioral principles, and is the one that was discovered first. In short, respondent conditioning is when we learn through association. The most famous example of this principle is the conditioning of dogs to respond by salivating when hearing the sound of a bell, an experiment carried out by Ivan Pavlov. How does this work?

All organisms are born with a set of reflexes. These are automatic responses to certain stimuli. For example, if we get a puff of air blown into our eyes, we will blink. In this case, the puff of air in the eye is the stimulus, and the blinking is the response. When we taste food, we begin to salivate. In this case, the taste of food is the stimulus, and the response is the salivation. For inborn reflexes like this, the stimulus and response are called "unconditioned stimulus" (UCS) and "unconditioned response" (UCR), respectively. What this means is that the relation between the stimulus and response in this case has not been learnt, but exists prior to any conditioning.

An unconditioned stimulus can be paired with a stimulus that does not have an inborn relation to the unconditioned stimulus. This is what happens in the famous example with the dogs salivating to the sound of the bell. We begin with an unconditioned stimulus-response relation: the taste of food is an

unconditioned stimulus for the dogs, that leads to the unconditioned response of salivating. Then we can take another stimulus, the sound of a bell, which does not have an inborn relation such that it automatically results in the response of salivating. For this reason, it is called a “neutral stimulus” (NS). If the sound of a bell is rather consistently followed by the delivery of food that the dog can taste, then the neutral stimulus, NS, (the sound of the bell) becomes paired to the unconditioned stimulus, UCS, (the taste of the food). This happens because the neutral stimulus begins to be a predictor of when the unconditioned stimulus is about to occur. When this pairing has been established, the formerly neutral stimulus now automatically triggers the response of salivating. When this has occurred, we call the sound of a bell a “conditioned stimulus” (CS) instead of neutral stimulus, and the behavior of salivating in response to this conditioned stimulus is in this context called a “conditioned response” (CR).

Expressed in a shorter format:
UCS (food) -> UCR (salivation)
NS (bell)
-- --
CS (bell) -> CR (salivation)

These are the basics of respondent conditioning. Next, let’s look at operant conditioning.

2.2.2 Operant conditioning
Operant conditioning (also called instrumental conditioning) was the second behavioral principle to be discovered. In short, operant conditioning is when we learn through consequences. A famous example is a rat pressing a lever in order to get food. This principle is a little more complex. And while respondent conditioning exists for practically all organisms, operant conditioning exists only in organisms that have a certain degree of complexity. So how does this work?

Operant conditioning can be described as involving three main components: a behavior (B), an antecedent (A) which is a stimulus that comes before the behavior, and a consequent (C) which is a stimulus that comes after the behavior. Diagrammatically, this can be laid out as follows in the temporal order:
A - B - C
The antecedent can be literally any type of stimulus, it doesn’t have to be an unconditioned or conditioned stimulus. Likewise, the behavior can be any type of behavior. The consequent, however, must be of a certain type in order for operant conditioning to occur. The consequent stimulus must be either a reinforcer or a punisher. A reinforcer is a type of stimulus that increases the probability of the behavior (B) being performed in the presence of the antecedent stimulus (A). While a punisher is a type of stimulus that decreases the probability of the behavior (B) being performed in the presence of antecedent stimulus (A). Outside of behaviorism, reinforcers are also called rewards, and are described as appetitive to the organism. And punishers are described as aversive to the organism. In behaviorism, these properties are not important however, because one is only interested in the behavior and how it changes in relation to stimuli. Worth noting is that the behavior that has either been reinforced or punished in operant conditioning is called an *operant*, a term we will be using in the rest of the essay. An operant is a behavior that is performed because of an expected consequence.

Let's look at an example of reinforcement to see how this works. In the example of the rat mentioned above, we have a situation where a green light appears (an antecedent stimulus A), the rat then presses a lever (the behavior B), and this behavior is followed by food being released for the rat to eat (the consequent C), which has a reinforcing effect on the rat's behavior. There are different notations for these different terms. For simplicity I will use A, B and C, and indicate whether the consequent is a reinforcer or punisher by adding a +sign or -sign after the C.

In short form:
A (green light appears) - B (presses lever) - C+ (food is released)

With more repetitions of this sequence, the probability of the rat pressing a lever when the green light appears increases. When the operant conditioning has occurred, the behavior is performed in response to the antecedent stimulus *because* it is expected to lead to the reinforcing stimulus. In other words, it is done with a purpose of achieving something in the near future. In contrast, the conditioned behavior in respondent conditioning is automatically elicited.

There are more complexities involved regarding these principles, but they lie outside the scope of this essay. The interested reader may be referred elsewhere (e.g. Catania, 2007)[1]. However, there are some aspects of respondent and operant conditioning that may be necessary for our purposes: generalization and discrimination. In particular, these aspects are important for understanding how a behaviorist theory of language is possible, as we will see later when we look into Relational Frame Theory.

2.3 Features of respondent and operant conditioning: Generalization and Discrimination

An important aspect of respondent and operant conditioning is the fact that no stimulus is exactly the same as another. For example, the sound of a bell will never be exactly the same on different occasions. Sometimes it may be longer, sometimes shorter. Sometimes it will be higher in pitch, sometimes lower. If respondent and operant behavior only worked for stimuli that were exactly the same as the stimuli during the actual conditioning situation, this would make the conditioning much less useful. But we could also have the opposite problem if the respondent and operant behavior is elicited by a too broad set of stimuli that only had some remote similarity to the stimuli it was initially conditioned to. The processes of generalization and discrimination regulate this for optimal responding.

A stimulus can be generalized such that the same behavior is performed in the presence of an ever greater variety of different stimuli. We can also have the opposite change in sensitivity to stimuli, i.e. that the response is elicited by a narrower set or variation in stimuli. This is called discrimination. This can occur for stimuli in both respondent and operant conditioning. Also the operant *behavior* can be generalized and discriminated. This means that in the presence of the same antecedent stimulus, different *behaviors* may be performed as long as they serve the same function.[2]

[1] And for a more general introduction to these principles see Törneke (2010)
[2] Note that all stimuli and behaviors are "generalized" to some degree. The term "generalized" is used to indicate a more broad or diverse generalization than what is normally the case.

These features are of great importance in understanding how behavioral principles can explain human language and thinking, especially that of operant generalization, because linguistic or verbal behavior is a specific type of generalized operant, as we will see in section 2.6. These features are related to questions of how we can use the same word (e.g. “chair”) for stimuli that have very different physical features (e.g. a black chair with short legs, and a red chair with long legs), and how we can apply words like “is bigger than” to things that we have never before applied such a comparison to.

2.4 The basic terms of functional contextualism: Function and context

Now that we have some understanding of the basic terms and principles of behaviorism, we can begin to look at the two features that are of central importance in the CBS philosophy functional contextualism: function and context. We can then turn to the more intricate details of this philosophy.

2.4.1 Function

One of the major features of the philosophy of functional contextualism is that of function. This involves both function of behaviors and function of stimuli.

2.4.1.1 Function of behaviors

As was mentioned above, the operant serves a certain function, it is directed towards a goal. When a behavior is an operant, that behavior is performed because of the predicted consequent stimulus that follows. Respondent behavior on the other hand is reflexive, and is not performed in order to acquire a goal. The respondent behavior will be performed regardless of the consequences that follow that behavior. Thus, only operant behavior has a function.

In fact, the operant is even *defined* by its function in CBS. This means that in behaviorist terms, two different behaviors are said to be the same if they serve the same function. For example, the behaviors of pressing a light switch with my hand, and pressing the light switch by jumping high in the air and pressing it with my foot, may be seen as very different if we look at the physical features, or *topography*, of these behaviors (i.e. their movements). But if I perform these behaviors for the same reason, e.g. to switch on the light, then they serve the same function and belong to what is called the same *functional class*. This is contrasted with what’s called the topographical definition of a behavior, where we say that two behaviors are the same if they have the same physical features, regardless of what function they served. Two different behaviors of hitting my dog on the head is seen as the same behavior topographically, even if in one case I hit my dog on the head because he had just swallowed something dangerous and I need to get it out of him to help him survive, and in the other case I hit my dog because I enjoy hurting other beings. In this case, the two behaviors are seen as different functionally, because they seek to achieve different goals.

This is a major difference between functional contextualism and descriptive contextualism, as mentioned above, and is a core reason for the pragmatic stance of functional contextualism. It is also something that separates functional contextualism from the other worldviews of mechanism, formism and organicism, where focus may be more on mere description of events.

2.4.1.2 Function of stimuli (Stimulus function)

We can also talk about function of stimuli. In this case, however, the functions of a stimulus are those responses which a *stimulus* elicits. So a function of the unconditioned stimulus “food” (or the

conditioned stimulus “bell”), is to elicit the unconditioned response “salivate”. A function of the antecedent stimulus “green light” is to elicit the operant response “press the lever”.

But stimulus function is more than this, it also includes other types of responses, namely (perceptual) responses of seeing, hearing, smelling, tasting, feeling the touch of. So when a car appears in front of me, this stimulus elicits the responses of seeing the shape of the car, seeing its colour, hearing its engine, maybe smelling the exhaust, and if I touch it also feeling the smoothness of the metal exterior. In the contextual behavioral paradigm, these are responses too. What this means is that what is usually called “properties” of objects and events in philosophy, including when discussing the correspondence theory of truth, can be described as stimulus functions in CBS. This will be of importance when we attempt to make sense of the correspondence theory in CBS terms.

2.4.2 Context

Next, we turn to context. According to functional contextualism, context is not just a curious aspect that we can choose to look at for some additional information about a stimulus or behavior. Quite the opposite. According to functional contextualism, nothing can be understood properly about stimuli and responses except in the light of their context. There are different types of contexts, grouped into historical context and current context.

2.4.2.1 Current context

The current context are those features that appear or are present in the temporal and spatial vicinity of the responses under investigation. Since the conditioned stimulus in respondent conditioning and the antecedent & consequent stimuli in operant conditioning are part of the context of the conditioned responses, we can see from the nature of these principles why the only way to make sense of the behaviors are by taking this context into account.

2.4.2.2 Historical context

Historical context refers to the history of the organism. In particular, this includes the *learning history* of the organism. The reason that a stimulus is a conditioned (as opposed to a neutral) stimulus to an organism is because of the events that have occurred in such a way to allow that to happen. Similarly, an operant behavior is performed because of the previous history of that behavior being reinforced in the presence of the antecedent stimulus on several occasions. The learning history as a concept is something that stands in contrast to the concept of memory, which is a concept more readily used by non-behavioral psychology, or cognitive science. CBS deals with what is publicly or privately observable. A memory in the sense of some hidden structure that causes behavior (including experiences of past events), is not observable but can only be inferred.

This view of previous events being seen as part of the historical context can be extended to incorporate the *evolutionary history*. This is the history of the species, in contrast to the specific individual. It is this evolutionary history that is responsible for the different principles and unconditioned stimulus-response relations that the organism did not learn throughout its lifetime, but where instead born with. For example, the reinforcing effect of food is the result of it having had an evolutionary function in helping the species survive.

2.4.3 The unit of analysis of functional contextualism

It is for these reasons that the basic unit of functional contextualism is not just the act (or behavior, or response) itself, but the act-*in-context*. It is written with hyphens to illustrate the inseparability of the parts. What this means is that one is not interested in the behavior in its isolation, but only to what function (or purpose) that this behavior serves in the specific context, noting that functions will differ depending on the context.

I believe that these philosophical principles of function and context are not mere assumptions or the result of some mere preference. Function and context are necessitated by the empirical results. A proper analysis of a behavior cannot be done without taking these into consideration, at least if we want to be able to understand the meaning of or reason for a behavior, and not just some mere description of its physical features. This is why these concepts are so central to functional contextualism. But on top of these principles there are also some further philosophical principles, that may be thought of more as assumptions and the result of preference, which is explicitly admitted in the CBS- literature (e.g. Hayes, 1993). Let us now turn to what these are, and how the different philosophical principles relate.

2.5 Functional contextualism and pragmatic truth

Functional contextualism is an attempt at making a philosophy of behaviorism that is explicit and coherent, in contrast to radical behaviorism that is believed to be inconsistent. This has boiled down to formulating a set of components seen as necessary and sufficient for a complete and coherent philosophy of psychological science: the unit of analysis (act-in-context), the goal of analysis (prediction-and-influence with precision, scope and depth), and the view on truth (pragmatic truth criterion) and ontology (a-ontological stance). For an analysis to be CBS-compatible, it must be in agreement with these principles. Thus, in order to be able to see if a correspondence theory of truth may be compatible with CBS, we have to take a look at these. Let's begin with the analytic goal of functional contextualism.

2.5.1 The analytic goal

The goal of analysis in functional contextualism is *prediction-and-influence with precision, scope and depth*. It is important to connect this with the unit of analysis of functional contextualism that was presented earlier: the act-in-context. This means that the goal applies only to analysis of behaviors, or more specifically, behavior that is defined in terms of its function in a specified context. The goal is therefore sometimes expressed as prediction-and-influence *of behavior* with precision, scope and depth (Levin, Twohig & Smith, 2016). Another variant of how the goal has been expressed is with the words "based on verifiable experience" added at the end (Hayes, Barnes-Holmes & Wilson, 2012), to emphasize the philosophical principles of observability as we talked about in part 1. For simplicity I will stick with the first formulation. Let's now look at what this goal means more precisely.

Prediction-and-influence

Prediction-and-influence is a joint goal. In other words, CBS is not interested in mere prediction. According to CBS, science should help us shape our world, and thus mere prediction is not a useful goal. If we look at correlational data we might be able to see that a person is likely to fail an upcoming test, but that doesn't mean we also know how to stop it from happening. If we on the other hand look at how manipulations of different factors affect the person's ability on tests, then we will be able to both predict and influence such events. This is why functional contextualism has prediction and influence as a combined goal, and why it is written prediction-and-influence. This differs from a lot of

cognitive research, where mere prediction, or even mere *description*, of events may sometimes be seen as a sufficient goal (Hayes, Barnes-Holmes & Wilson, 2012).

The goal of prediction-and-influence of a behavior is said to be achieved when an analysis has identified the variables that permitted the prediction of the behavior, and these variables would, if manipulated, influence the probability of the behavior occurring (Biglan & Hayes, 1996). This also relates to the principle of "verifiable experience" or observability. We cannot directly influence hypothetical constructs, nor can we directly observe the effect on hypothetical constructs. This means that they are not useful for the goal of prediction-and-influence.

While the goal of "prediction-and-influence" refers to the *statements* (i.e. the verbal analysis) being made about a behavior, "precision, scope and depth" refers to the *concepts* being used in the analysis or statement, not to the statement itself.

Precision, scope and depth

In functional contextualism, *precision* means that the concepts used when analyzing a behavior are unambiguous and do not overlap. This has the result that a limited number of concepts are needed to explain a behavior in a certain context. For example, the concept of "conditioning" is thought of as having great precision because it is unambiguous and clearly delineated from other behaviorist concepts. Precision does not refer to the idea that a prediction-or-influence of a behavior in a certain situation is very accurate, or more accurate than other predictions or influence. Rather, behaviorism considers itself to have found concepts that *generally* allow for analysis that does not involve superfluous concepts. In other words, it's closely related to the idea of parsimony of concepts. Next, *scope* means that a given concept applies to a range of cases, such as behaviors of animals as well as humans, in understanding behaviors such as studying as well as socializing. Terms like "response" and "stimulus" are so general that they apply in any type of situation where we are analyzing a behavior. Lastly, *depth* means that the concepts used in an analysis are consistent with other fields, or levels of analysis. The clearest examples of such concepts are "variation" and "selection", which have recently started to be used more in CBS, which apply both to the analysis of biological, cultural and individual evolution. (Hayes, Barnes-Holmes & Wilson, 2012). But all behaviorist concepts are believed to fulfill all three of these goals.

This means that for an analysis to be CBS-compatible, it must use terms that have precision, scope and depth, and it should allow for prediction-and-influence of the behavior under analysis. Let us now look at the pragmatic truth criterion, and how it relates to this goal.

2.5.2 The pragmatic truth criterion

According to functional contextualism the truth criterion is that of "successful working", i.e. that something is true if it works to successfully accomplish a goal. But not just any goal, namely the analytic goal of functional contextualism. Also, it is argued (although on dubious grounds) that the pragmatic truth criterion applies only to verbal statements (Hayes, 1993). This would yield the following definition:

> **Pragmatic truth criterion:** a statement about a behavior (act-in-context) is said to be true to the extent that it leads to the achievement of prediction-and-influence of that behavior, using concepts that allow for precision, scope and depth.

This leads to some problems, however. First of all, one would expect a theory of truth to apply to more than just statements about behaviors, which does not seem to be the case. For example, what does this mean for evaluating the truth of a statement like “God exists”? It seems to me that we have two options. 1) we treat the statement “God exists” (or rather, the act of expressing that statement) as a behavior that we analyse by looking at the purpose with which it is expressed in the specified context (e.g. to give comfort), and then express our analysis in verbal terms with the goal of achieving prediction-and-influence of that behavior. In this case our *analysis* of the statement will be true if the goal of prediction-and-influence with precision, scope and depth is achieved, but it will be left unsaid whether the original statement *itself* is true or not. Or 2) we treat the statement “God exists” as an instance of analysis, in which case the statement will be true if it is a verbal analysis of a behavior and achieves the goal of prediction-and-influence of that behavior with precision, scope and depth. But it does not seem to be a verbal analysis of a behavior, which means that the pragmatic truth criterion does not apply to this statement, and can say nothing of its truth. The same reasoning would apply to statements like "Snow is white", a paradigmatic example of a statement that the correspondence theory of truth applies to. So the pragmatic truth criterion would be unable to say anything about the truth of statements that we are generally interested in knowing the truth about.

But there is yet another problem. In the CBS literature, the pragmatic truth criterion is not neatly summarized as I did above, but rather is expressed in separate pieces. Also, almost no examples of pragmatic truth are given in the CBS literature, and the only example I have come across seems to contradict their own definition of the pragmatic truth criterion. I quote:

> ‘Consider for example two different renderings of a building: one is an artistic drawing of the building in perspective, and the other is a blueprint of the building. Which is the “true drawing” of the building? Both are “representations” and the contextual approach would hold that there is no “true drawing” in any objective sense. The truer drawing could be determined only in the context of the specific goals or purposes that apply. If one needed a drawing in order to identify the building while walking down the street, the perspective drawing would be the more useful and thus “truer” - in the sense that it is true for this purpose. Alternatively, if we wanted to know how to safely remodel the building, the blueprint would probably be a truer representation.’ (Hayes, Strosahl & Wilson, 2011 pp.33-34)

This quote is problematic for a number of reasons. Firstly, the stimulus being evaluated is not a statement (especially not a statement about a behavior), but a drawing, and thus cannot be an instance of pragmatic truth, according to the theory. Secondly, the example uses the term “representation”, which is a term that CBS considers problematic because of its mechanistic implications. The fact that the word is placed in quotes does not change the fact that they appear to need the term to explain what they mean. Thirdly, the example states that there is no truth in an “objective sense”, which would imply that pragmatic truth cannot be agreed upon by different people but is just a relative concept. I do believe, however, that this is just a sloppy way of writing that pragmatic truth is sensitive to the context. Fourthly, the example puts “truth” or “true” in quotes (as is common in the CBS literature) which suggests that they do not consider it to actually be truth they are talking about, but are using the term in a more metaphorical sense. This point is reinforced by another quote. “All therapeutic interactions are evaluated as they relate to the client’s chosen values and goals, and the issue is always workability - that is, whether they work in practice - not objective truth.” (Hayes, Strosahl & Wilson, 2011). This passage suggests that the issue with correspondence or objective truth is not that there is

no such thing, or that it does not make any sense, but rather that truth isn't important, only workability is. That puts into question whether workability (or successful working) despite the previous reasoning really is best considered as a theory of *truth*, or whether it would be more fitting to treat it as just a goal, a goal that is considered more valuable than the goal of finding the truth.

For these reasons, it is unclear whether the pragmatic truth criterion can really be considered a valid truth criterion at all. It seems to work just as well for all practical purposes of CBS to simply say that a goal of analysis is that it should be useful for the purposes of achieving our goals, or more specifically, the goal of prediction-and-influence with precision, scope and depth. Thus, there may be an opening for the correspondence theory of truth. Also, when investigating whether a theory or analysis is CBS-compatible, I will take this to not include being compatible with the view that truth should be defined in this pragmatic way, but simply that it has to be compatible with the other components of functional contextualism. Next, we turn to the ontological stance of functional contextualism.

2.5.3 The ontological stance

Functional contextualism takes what's called an *a-ontological stance*. What this means is that it is principally disinterested in questions of ontology, i.e. about the real nature of the world. I will let the following quote illustrate this position:

> Issues of epistemology (how we know what we know) become the core focus of functional contextual philosophy of science, while issues of ontology (whether what we know is real and what the real categories are) become uninteresting or irrelevant. Claiming that something works "because it is real" adds nothing to workability. Since there is no difference that does not make a difference, questions of ontology are simply put aside within a functional contextual approach. This is not due to idealism or dualism but rather reflects the practical imperative of a contextualistic approach. (Biglan & Hayes, 2016)

In other words, it doesn't have any use for the purpose of prediction-and-influence with precision, scope and depth to talk about the ontological status of things. One interpretation of this is that whether someone talks about reality or not doesn't really matter, that it's okay either way. But in practice this stance involves an *objection* to talk of "reality". More specifically, it is an objection against statements that imply that talk of "reality" does have some explanatory power that according to CBS it does not have.

But exactly what is meant by the term "real" is not entirely made clear. In some respects it seems close to the Kantian sense that the objective reality (Ding-an-sich) can never be reached, and therefore any attempts at establishing the reality of things is doomed to fail (see for example Kant's (1781 [1998]) Critique of Pure Reason). Or at least that we wouldn't be able to know if we did find the objectively true description of reality. And even if it did find it, it still wouldn't matter for any behaviorist purposes. This view is connected to the fact that all descriptions are made using language, and language is always an act performed by a human, thus affected by the context and function of this act. The point is formulated in the following quote:

> "Language itself begins and ends as nothing but a social behavioral tool, not a passageway to pre-organized reality" (Hayes, Barnes-Holmes & Wilson, 2012).

Here, “pre-organized reality” is supposed to be understood as the (by nature) pre-given concepts or categories, or descriptions. Thus, the ontological objection of CBS is directed towards the idea that there is a single description that is considered to capture the true nature of reality best. This view also seems related to the idea of quantifier variance by Eli Hirsch (2011) which is that there are several (ontological) languages i.e. several ways of dividing up the world, that reflect reality equally well.[3] But while compatible with this view, CBS has chosen to avoid all talk of reality, instead of accepting that a plurality of descriptions may all be considered to reflect reality equally well, but serve different purposes. The scepticism towards ontology should, however, be distinguished from another usage of the term “ontology” as outlined in the following quote:

> Some use “ontology’ to refer merely to an explicit specification of concepts to be used in a domain, and the relationship between them. If that is all that is meant by the term, there is no necessary conflict with contextualism. If, however, one means the more traditional philosophical definition of what categories exist or can be said to exist in the world and the correspondence between analyses of the world and these categories, then there is a major conflict caused by the decontextualized truth criterion being employed. (Biglan & Hayes, 2016)

Another aspect of the a-ontological stance concerns objections to ontological commitments regarding making events into some kind of entities (an action called “reifying” - making real). For example, treating “relation” as an abstract form of entity, or the act of remembering as an instance of the entity “memory”. In the CBS literature there is a frequent idea that if some phenomenon is expressed as a noun, then this would imply that it is to be understood as an entity and involves ontological commitment. For this reason, when talking about relational frames, this is sometimes rewritten as “relational framing”, to clarify that they don’t consider relational frames as abstract entities, but as behaviors or events. But why the use of a noun would imply “entity” or ontological commitment is unclear. I believe this position can clearly be invalidated by the fact that the words “behavior”, “event”, “context” and “function”, the most central terms of CBS, are all nouns, but it would be preposterous to claim that this means that these are all entities. Lastly, it also remains unclear why it could never be useful for some purpose to imply the existence of abstract entities. Talk about “memory” as some abstract entity seems useful at least in some contexts.

For clarity, in this essay I will use the following definition of the a-ontological stance:

> **The a-ontological stance**: it is not acceptable to imply the existence of abstract entities or a correspondence to a pre-organized reality

This stance is not the same as the less nuanced stance that no use of words relating to “reality” is acceptable, *regardless* of what they actually imply in the present context. The above definition is what will be used when attempting to formulate a CBS-compatible correspondence theory of truth in part 4.

2.5.4 Summarizing the philosophical view of CBS

As we have seen, the functional contextual philosophy has implications for what is considered valid objects of analysis, as well as how this analysis is to be done according to CBS. This includes a CBS

[3] called “quantifier variance” because the theory is expressed in terms of quantifiers like “there is” or “there exists” possibly having different meanings in different contexts.

analysis of the correspondence theory of truth, which we will look at in part 4. These implications can be summarized as follows.

For an analysis to be CBS-compatible:
1. That which is analyzed must be (treated as) a response/behavior
2. This response, as well as the surrounding stimuli and responses, must be understood (defined) in terms of the functions they have in the specific context
3. The analysis may only be done using CBS-terms (that would allow for prediction-and-influence with precision, scope and depth)
4. The stimuli and responses must be (at least in principle) observable (either publicly or privately), and cannot be hypothesized unobservable/abstract entities
5. Our analysis cannot be claimed to correspond to a pre-organized reality

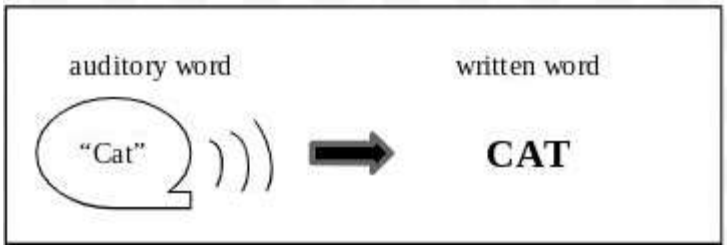
But in order to talk about statements/analyses in CBS-terms, we will need to look at the CBS theory of language: Relational Frame Theory. This will allow us to determine whether the correspondence theory actually is CBS-compatible, despite the CBS-objections against such a theory.

2.6 Relational Frame Theory

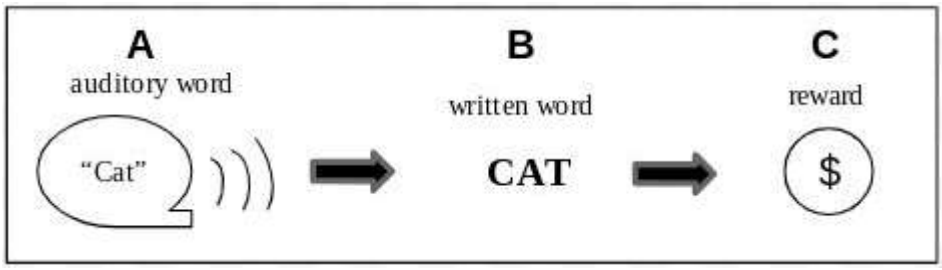
Relational Frame Theory (RFT) is the latest behavioral principle or theory. In short, RFT explains a behavior only seen among humans, namely that of language, and all those things that have to do with language. This includes talking, thinking, writing, using symbols, doing mathematics. With this principle we take yet another step in degree of complexity. While practically all organisms are capable of respondent conditioning, only more sophisticated organisms such as animals are capable of operant conditioning, and lastly only humans are capable of relational framing. So how does this work?

2.6.1 Stimulus equivalence

The basis of (or seed for) relational framing is a phenomenon called stimulus equivalence. It was discovered in 1971 by the behaviorist researcher Murray Sidman (1971). The seminal study had one participant: a 17-year old boy with intellectual disability, who did not know how to read. In the experiment he was taught to select the matching visual (written) words in response to auditory (spoken) words. For example, when he heard the word ‘cat’ he was supposed to select the written word ‘cat’, when he heard the word ‘dog’ he was supposed to select the written word ‘dog’.

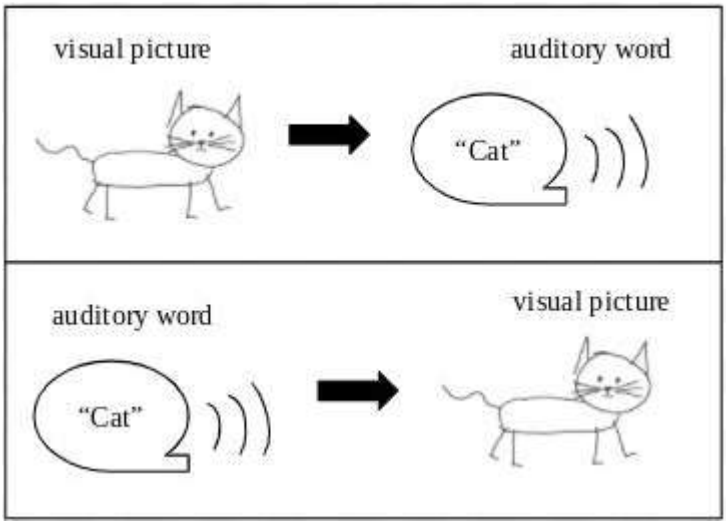


For each correct selection, the boy was rewarded with money and sweets. Learning the matching between an auditory and a visual word is thus an example of operant conditioning. The auditory (spoken) word is the antecedent stimulus (A), selecting the correct visual (written) word is the behavior (B), and the reward of money and sweets is the reinforcing consequent (C+).

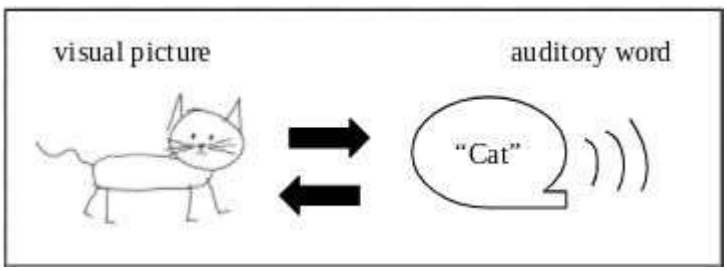


With enough training, this kind of operant conditioning will lead to the behavior of selecting the matching written word 'cat' to be performed reliably in the presence of the auditory or spoken word 'cat'. This is also what happened for the boy in this experiment. Similarly he was able to do the same thing for the other words in the experiment. No surprise there. The interesting part is what came next.

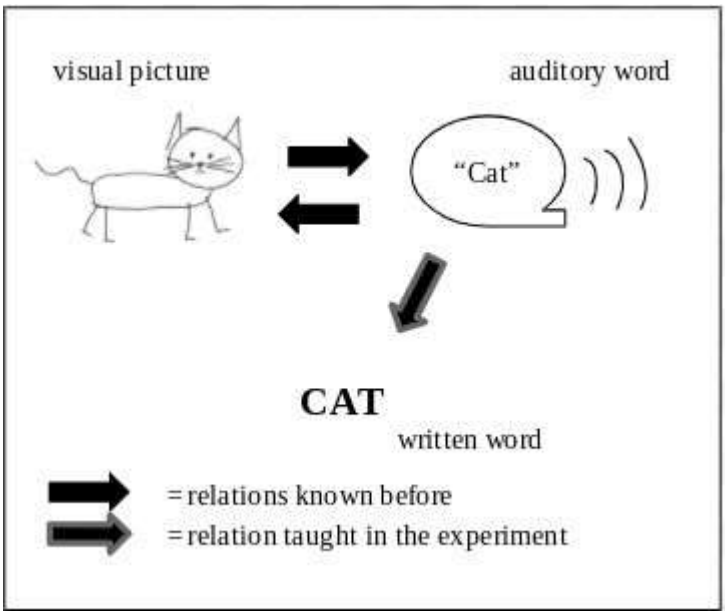
Although the boy did not know how to read before the experiment, he did know the names of the animals and objects involved in the experiment. In other words, upon hearing the word 'cat' he was able to select a picture of a cat, and when shown a picture of a cat he was able to say the word 'cat'.



Given what we know about operant conditioning, we can conclude that these abilities too must have been the result of operant conditioning. In the first case, when the boy had heard the spoken word 'cat', his behavior of selecting a picture of a cat must have been reinforced in some way, for example by being praised by one of his caretakers (which is a primary or unconditioned reinforcer). Similarly, the behavior of saying the word 'cat' when shown a picture of a cat must also have been reinforced. Hence, the boy had learnt to produce or indicate one stimulus when presented with the other, and vice versa in the other direction. We can express this bidirectionality in the following way:



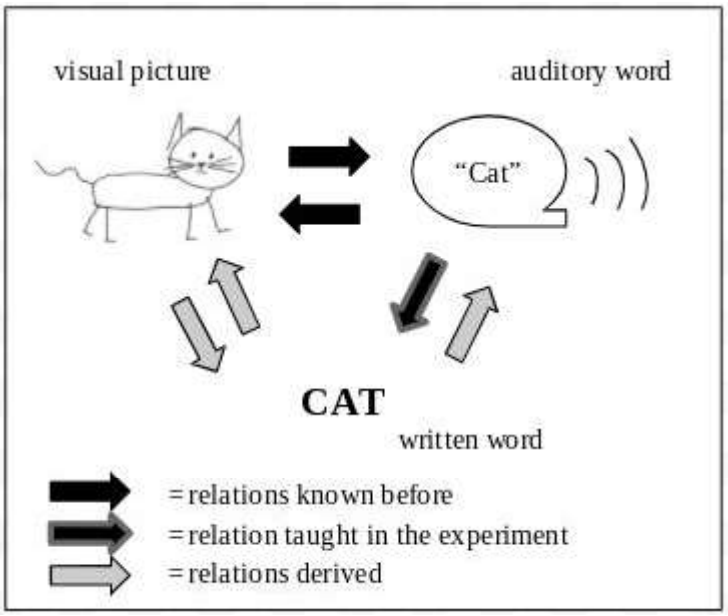
Consequently, after having been taught the new relation described above, the boy had in total been taught the following relations between stimuli:



The interesting thing was when the boy was subsequently shown visual words, and asked to select the matching picture. Looking above at the relations that he had been taught, there had never been any training of responding to a visual word with selecting a visual picture, or even selecting (or speaking) an auditory word in response to a visual word. The visual word had never occurred as an antecedent stimulus in an operant conditioning, and thus the behavior following this stimulus should be completely random according to the principles of operant conditioning. Instead, the boy was able to match the correct picture quite perfectly.

This was a *big* deal. According to the principles of respondent and operant conditioning, an organism should only be able to display evidence of having learnt a certain stimulus-response relation if this relation had been explicitly taught. But here was clear evidence that the boy had learnt something that he had not been explicitly taught, rather the new behavior had been *derived* from previous learning.

The behavior of selecting the correct picture when shown a visual word was not the only stimulus-response relation that the boy had derived. When the boy was subsequently presented with a picture first, he was able to select the correct written word too, despite not having been taught to do so. He was also able to say the spoken word when presented with the written word, despite not having been taught to do so. Adding up all the displayed stimulus-response relations that the boy displayed we get the following relations:



The way that Sidman made sense of this was to say that the boy had learnt that the three different types of stimuli were equivalent to each other. Thus, since the boy had learnt that pointing to a certain picture when hearing a certain word would be reinforced, he now had concluded that if the auditory word was replaced by the written word, then pointing to the same picture would still be reinforced, because the spoken and written words are equivalent. And so on, for all possible substitutions between stimuli.

Given the nature of the above experiment, i.e. its ability to show how we learn the relations between linguistic stimuli (i.e. spoken and written words) and objects (or pictures of objects), it is natural to assume that the phenomenon of stimulus equivalence may be the missing key in explaining human language in behaviorist terms.

Let's begin with looking at how to make sense of stimulus equivalence from the perspective of RFT.

2.6.2 The relational frame[4]

According to RFT, stimulus equivalence is an example of a *relational frame*. What characterizes a relational frame? The term "relational frame" or "relational framing" is a shorthand for the longer term "arbitrarily applicable relational responding". What does this mean? Well, let's begin by looking at the last part of that expression: relational responding.

Relational responding means that we respond (or behave) in accordance not with a stimulus in itself, but in accordance with a relation that a stimulus has to other stimuli. A simple example is if I am at a café and am interested in having a cinnamon bun, I may choose the *biggest* cinnamon bun that I see. In other words, my behavior of picking this cinnamon bun is not controlled by the properties of that bun in itself, but by the relation it has to the other cinnamon buns in its vicinity. If a new cinnamon bun is placed among them as I stand there and watch, and this bun happens to be even bigger, I will decide to choose that one instead. All animals display relational responding. It is possible to teach a dog to bring me the biggest stick in the park if I reinforce its behavior everytime it does so, but not at times it does not. Similarly, I could teach it to bring me the smallest. If the dog always had to choose

[4] For a general introduction to Relational Frame Theory, see Törneke (2010)

among three sticks I could even teach it to always bring me the medium sized one. But all of these forms of relational responding are based on the physical characteristics of the stimuli involved, or as it is called in RFT: on the *formal properties* of the stimuli involved, even if the response is based on the relation between these formal properties.

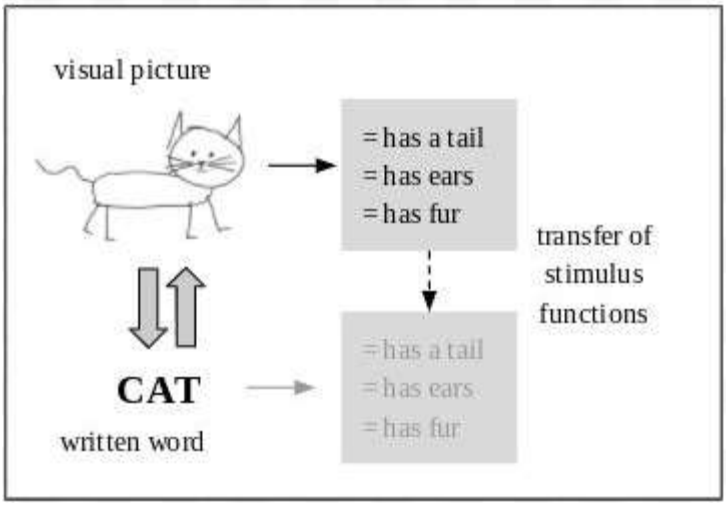
Relational framing, on the other hand, is a different type of relational responding, namely an *arbitrarily applicable* relational responding (or AARR). This means that the relational responding is not based on the formal properties of the stimuli involved, but rather on completely arbitrary relations between stimuli, such as those in the experiment of stimulus equivalence. The auditory stimulus 'cat' has no formal properties in common with the visual stimulus of a picture of a cat, or of an actual cat. They do not look the same, do not sound the same, do not smell the same, do not taste the same, do not feel the same. Yet, as the experiment showed, the stimuli were treated as the same, or equivalent. When the boy responded to the written word by indicating the matching picture, he did so because not because the written word and the picture looked the same, no he responded to the written word in terms of its relation to the spoken word, with which it was in a symbolic relation with (i.e. stimulus equivalence). This is the true essence of language. We can take any stimulus and use it as a symbol for something else. I can say that "blorgh" is just another word for "book", and if I in the rest of this text use the word "blorgh", you will understand that I am talking about books. The nature of the stimulus does not matter, it is completely arbitrary which stimulus I choose.

In RFT, stimulus equivalence is said to be an example of a relational frame of coordination. The frame of coordination is a broad category of different relations that involve examples such as "is the same as", "is like", "is similar to". RFT has also established other types of relational frames, such as relational frame of opposition, or relational frame of hierarchy. In contrast with frames of coordination, a frame of opposition establishes a relation between stimuli that treats them as opposites, not equivalents (e.g. white is the opposite of black). A frame of hierarchy establishes a relation between stimuli as one stimulus belonging to another (e.g. a car is a type of vehicle). These categories are very broad and are used in the research to establish the empirical features of basic relational types. But I don't think that this is supposed to be understood as any form of objection against more narrow categories of relations, like "has the colour" or "is the father of". It is simply that for the use of the empirical science, the broader categories are more useful for expanding the theory. In the rest of this essay I will therefore take any type of relation between stimuli to be an instance of a relational frame, and that the properties of relational frames apply.

2.6.3 Mutual entailment, combinatorial entailment, and transformation of stimulus function
For a relation to be classified as a relational frame it must exhibit three features: mutual entailment, combinatorial entailment and transformation of function.

Mutual entailment means that if stimulus 1 is related to stimulus 2 in a certain way, then the corresponding relation in the other direction is derived. E.g. learning that if the sound "cat" is the same as the written word "cat", then the written word "cat" is the same as the sound "cat" (the grey arrow on the right side in the diagram before). Combinatorial entailment means that if relational frames are applied between three or more stimuli, then the mutual entailment between the stimuli combine (the two grey arrows on the left side in the diagram before).

Transformation of stimulus function means that if a stimulus enters a relational frame with another, then the stimulus functions may be transferred or transformed in accordance with the relational frame. This is the feature that will be most important when we later apply RFT to the question of truth. So for example, if the stimulus “cat” enters into a frame of coordination with a picture of a cat, then the stimulus functions of the picture of the cat such as “has a tail”, “has ears”, “has fur”, may be transferred to the stimulus “cat”, such that the stimulus “cat” now also has these stimulus functions.



The examples of stimulus functions in the example above are formal properties of the stimulus that the written word “cat” is in a frame of coordination (stimulus equivalence) with. But as we learnt when talking about respondent conditioning, there are other types of stimulus functions. If a cat (or picture of a cat) has been conditioned such that it elicits fear, then such stimulus functions will transfer to the word “cat” as well.

Depending on the relational frame, the stimulus functions may not be merely transferred, but may be transformed. For example, if I learn that the stimulus “blob”, that I have no prior knowledge of, is the opposite of “ice”, then the stimulus functions may transform such that “blob” has stimulus functions such as “is hot” (instead of cold).

So far we have talked about single words. What about sentences? How can they be understood in terms of RFT? The RFT understanding of sentences involves what’s called relational contextual cues, of which there are two types.

2.6.4 Relational and functional context

Let’s take the sentence “Lisa is taller than Jens”. The terms “Lisa” and “Jens” are in stimulus equivalence with the respective persons, just like we have seen before. But the words “is taller than” are different. Together they form what is called a *relational* contextual cue, or Crel, which specifies how the stimuli “Lisa” and “Jens” are to be *relationally* framed. If I have met Jens, but not Lisa then the stimulus functions of Lisa’s height will be transformed in accordance with this relation such that I now expect Lisa to look taller than what I know Jens to be.

When a relational frame is established between two stimuli, it isn’t always clear which function is to be transferred or transformed. The job of the *functional* contextual cue, Cfunc, is to specify this *function*. For example, if a friend tells me about this new fruit he encountered over the weekend called Quap, and says “Quap is like lemon” it could mean that they have similar colour, similar shape, or

similar taste. The Crel "is like" does not specify which. But by adding "as regards to taste", or instead expressing it initially as "Quap *tastes* like lemon", then the word "taste" works as a Cfunc that specifies which stimulus function is to be transferred or transformed in accordance with the frame of coordination between the two stimuli.

Even if I can see the general point that is made when differentiating between the Crel and Cfunc, I believe these will be hard to differentiate in practice. It seems to me that Cfunc is really just a more detailed Crel. Therefore, from here on I will exclusively use the term Crel, and let it be an umbrella term that catches the meaning of both Crel and Cfunc as outlined above.

2.6.5 Implicit context

Crels are examples of explicit contextual cues, but there are also implicit contexts. We have a previous history of responding to different words and statements that affects the way we respond to them in the now. This historical context is rarely explicitly stated. For example, if someone says "Dogs are friendly", this is probably not meant in the sense that every single dog is friendly, but that most dogs are. We understand that because of the historical context of such statements. Similarly, a word may function to elicit stimulus functions while also serving a contextual function. For example, if quite formal words are used to describe something, we might infer that whoever talks about the topic is well-read on the research on the topic and therefore should be listened to. It would take too much time and energy to spell out all implicit contextual factors that it would not be helpful for the purpose of communication. But of course, in philosophy, such contextual factors can be of vital importance, because detail matters, and will be of importance in what follows.

I also want to mention that since relational frames are operants, this means that the aspects of function and context of stimuli in general apply to relational frames and linguistic stimuli, such that the same word can have different functions (or meanings) in different contexts, or that different words can have the same function (or meaning) in some contexts.

2.6.6 A new definition of language and cognition

So, according to RFT, language is defined as those behaviors that involve relational framing.[5] Since Contextual Behavioral Science views responses as including not only behavior observable by others, but also behavior only accessible to ourselves, that means that those aspects of our thinking that involve relational framing are also linguistic (or verbal) behavior. This means that cognition, in the sense of explicit thoughts that are internally observable, can be explained using RFT. There is no need for terms such as "mental representation", "processing", or "beliefs".

RFT is a young theory that is still under development. For further information about its development and how it may be able to answer the objections of Chomsky, the reader is referred elsewhere (e.g. Zettle et al., 2016). But there are some things that have not been explicitly specified in RFT that I believed can be argued to follow quite naturally from the basics of the theory, and that will help us in our endeavour to formulate a CBS-compatible correspondence theory of truth.

2.6.7 Some further postulations

[5] Since only humans can frame relationally, this directly excludes the behaviors of other animals as possible linguistic behavior, which was a big problem with Skinner's definition, which included any behavior that was trained by someone with linguistic knowledge.

2.6.7.1 Reference

Let's begin with the question of "reference". This is a term that is avoided in CBS literature. I think it is because of a fear that it will imply a mental world that refers to an external world, thus violating the a-ontological stance. I don't think such a fear is warranted, and that "reference" can actually be understood quite naturally in CBS terms. Let's take an example to illustrate this. If someone asks you "What does a cat sound like?", you do not answer by describing what the spoken words "a cat" sounds like, saying "it first has a short a-sound, then a hard k-sound, then an open aaa-sound, and finally a hard t-sound". Instead, you describe what cats (the animals) sound like: "They make a sound like 'meow'". Why do you do that? The spoken words "a cat" is in stimulus equivalence with the animal cat, and if a set of stimuli are stimulus equivalent, then it shouldn't matter which you use right? But it does matter. Because sometimes we are actually interested in one of the stimuli in stimulus equivalence, and not the others, and simply use one stimulus to talk about another. So how do we know which stimulus is referred to in different situations? One part of the answer is the Crel in the above statement, i.e. the words "sound like". This Crel means we can't be talking about the written words in the stimulus equivalence, because they don't have a sound. This is one example of implicit context.

But there is also a general implicit (historical) context constituted by our learning history of people asking us what we think about things using linguistic stimuli. When we have answered such questions by responding in terms of the object or event (as opposed to the written or spoken stimulus), our response has been reinforced. This means that there is an implicit context that makes us likely to treat written and spoken words primarily in terms of the objects or events that they are in stimulus equivalence with. Or to put it another way; we respond in terms of the stimulus functions that were transferred from these objects and events, and not of those from the other stimuli in the equivalence. Thus, it may seem like there is a unidirectional relation, where spoken and written words refer to objects and events, but not the other way around, but the relation is actually bidirectional, only that most contexts make use of only one direction. A simple example to show the bidirectionality is if someone points to a cat and asks "How do you spell that?". Since you have a learning history of the word "spell" applying only to written words, this time you will understand that the cat in front of you refers to the written stimulus "cat". Understood in this way, "reference" has nothing to do with a relation between a mental idea and an external reality, but is rather a relation between two equally observable stimuli, and therefore does not violate the a-ontological stance. I will therefore use this term in this sense in the rest of the essay.

2.6.7.2 Crels and stimulus equivalence

Crels were earlier distinguished from words that do not specify some relation, like "Lisa" or "cat". However, I think it is reasonable to conclude that Crels, just like other linguistic stimuli, are in stimulus equivalence with some other stimulus or behavior. What makes them different is not that they lack stimulus equivalence, but rather that they are in stimulus equivalence with something that is more of a relating kind. Thus, words like "is bigger than" are in stimulus equivalence with the perceived relative size between stimuli. One might perhaps worry that this would be some abstract entity of "comparison", but as long as we can *observe* something as being bigger than something else, there should be no objection on CBS grounds against claiming that a Crel like "is bigger than" is in stimulus equivalence with such an act of relating.[6] I believe this Crel may initially be tied to observed

[6] One may, however, wonder about the distinction between stimulus and response in such a case. Do the words refer to the act of observing, or to the resulting perceived relation between stimuli?

relative physical size, and then be generalized to include a broader set of stimuli. For example, we learn to apply this relational frame to celebrities, in terms of which celebrities are bigger (or more famous) than others.

A question related to this is what to make of linguistic stimuli that express a verb, but do not explicitly specify a relation between two stimuli, e.g. “to eat”. Nothing in RFT has been said about these so far, as far as I know. It seems that “eating” is in stimulus equivalence with the behavior of eating (or the stimulus that results from our observation of such behavior). One possibility is that such words function both to specify a stimulus, and a relation. If I say “John eats”, the word “eats” may thus both specify a stimulus (i.e. that of eating), and the relation that “this applies to the preceding stimulus”. This can perhaps be seen more clearly if we rewrite the statement as follows “John is in the act of eating”. This looks just like any other statement where we have a Crel that specifies a relation between two stimuli. So we can see that statements like “John eats” do express a relation between two stimuli, even if we perhaps didn’t see it immediately by simply looking at the words. Thus, some formulations may be equivalent in content, but differ in terms of the degree to which it makes certain aspects explicit. The present example shows one way of making relational frames more explicit, which may serve a purpose in certain contexts. I will make use of this when attempting a CBS-compatible correspondence theory of truth in part 4.

The fact that the relation between stimuli may be implicit is even more clear if we look at statements in Russian, where there aren’t even any words for verbs like “to be” or “to have”, where such relations are simply implicit. For example, “Dima medic”, or translated word for word into English: “Dima doctor”, states that the person called Dima is a doctor. The fact that this is understood as meaning that Dima is a doctor, and not that he hits a doctor, or that we are talking about doctors specialized in treating patients called Dima, is a result of the learning history of the speakers. Speakers of Russian know this relation without it having to be explicitly stated.

Now that we have looked closely at both the philosophy and the empirical science of CBS, it is finally time to return to the question of truth again, and discuss the correspondence theory of truth in more detail, before investigating the correspondence theory more closely from a CBS-perspective, and attempt to formulate a CBS-compatible version of this theory.

Part 3: The correspondence theory of truth

As has been explained previously, the CBS criticism against the correspondence theory of truth is based on it being seen as part of a mechanistic worldview. Because the Pepperian separation of worldviews into four categories may be overly generalized, this means that there is a risk that the correspondence theory as portrayed in CBS may be a mischaracterization of the correspondence theory as it has actually been proposed throughout the history of philosophy. This section will therefore outline the different types of versions of the correspondence theory as they have actually been proposed, before we begin to investigate whether this idea may be compatible with CBS in part 4.

3.1. The correspondence theory of truth in philosophy in general

What is truth according to the correspondence theory? A common short explanation is that truth is correspondence with reality, or a worldly fact. But due to the question of what “correspondence” is, and what is meant by “reality” or “a worldly fact”, there are different versions or formulations of the

correspondence theory, such that it is often described in the general terms of truth consisting of a relation (to be specified) to a portion of reality (to be specified) (David, 2016). Then there is of course the question of what kind of things we might ask of whether they stand in such a relation to that portion of reality. So in general, the correspondence theory can be said to consist of three parts, that together describe what it means for something to be true:

Truthbearer - Relation - Truthmaker

The truthbearer is that of which we are asking whether it is true or false. The truthmaker is that which makes the truthbearer true (and whose absence makes the truthbearer false). The relation is the way that the truthbearer relates to the truthmaker, in order to be true. The standard version of the correspondence theory is to take truthbearers to be statements or propositions, the relation to be correspondence[7], and the truthmakers to be facts or objects. This gives us the theory: a statement or proposition is true if it corresponds with a certain fact or object. For example, the statement "Snow is white", is true if it corresponds to the fact of snow being white.

The concepts used for the terms of truthbearer, relation and truthmaker vary greatly between different versions of the correspondence theory. The different versions can generally be categorized on the grounds of whether truth is described in metaphysical terms, semantic terms, or a combined version of the two, as well as a distinction between truth being described as object-based or fact-based (David, 2016). The different versions have their own advantages and disadvantages, and may be susceptible to different types of criticism.

The metaphysical version is that which characterizes truth as correspondence with reality, with emphasis on "reality". Often these theories take the truthbearer to be an *internal* idea or belief, to correspond with an *external* reality in the form of objects and events, differentiating between the knower and the known. But other metaphysical versions take the truthbearers to be more abstract like propositions. The semantic version, on the other hand, focuses on the relation of signification, and less on the nature of the truthmaker. This kind of formulation takes the truthbearer to be some kind of sentence, but is less explicit on what they take to be the truthmaker. A version of semantic correspondence theory may be expressed along the lines of "a sentence is true if it signifies the way it is". Then there is a combined version of both semantic and metaphysical components, where truthbearers are taken to be some kind of mental sentences, that signify something in the external reality.

As for the distinction between object-based and fact-based versions, object-based theories take objects to be truthmakers, while fact-based theories take facts to be truthmakers. Object-based theories can best be understood by using the example of a sentence or statement that has a subject term and a predicate e.g. "Snow is white", where 'snow' is the subject, and 'is white' is the predicate applied to the subject. A truthbearer (sentence/statement) is said to be true if and only if the predicate corresponds to a property of the object referred to by the subject (David, 2016). So "Snow is white" is true, if and only if the predicate "is white" corresponds to a property of the object that "snow" refers to (the visual 'object' "snow"). This version of the correspondence theory in effect involves more than one relation. On top of the *correspondence relation* between the predicate and the property of the

[7] Which of course is why this type of theory is called "correspondence theory".

object, it also involves the *reference relation* between the subject and the object it refers to. What is unclear is how this version applies to instances where the truthmakers don't have subject-predicate structure with corresponding objects. Fact-based theories, on the other hand, generally aren't considered to need truthbearers to have subject-predicate structure (David, 2016). This formulation is instead more abstract such as: "a belief is true when there is a corresponding fact, and is false when there is no corresponding fact" (Russell, 1912).

A list of the most widely used concepts that have been used as truthbearers, relations and truthmakers would include the following. For truthbearers: beliefs, thoughts, ideas, judgments, statements, assertions, utterances, sentences, and propositions. Relations: correspondence, conformity, congruence, agreement, accordance, copying, picturing, signification, representation, reference, and satisfaction. Truthmakers: facts, states of affairs, conditions, situations, events, objects, sequences of objects, sets, properties, and tropes. (David, 2016)

The combinations of the varying concepts above yield a vast number of different characterizations of the correspondence theory[8]. How does this compare with how the correspondence theory of truth is described in the CBS literature?

Given the CBS characterization of the correspondence theory of truth, as was presented in section 1.3, it seems clear that the CBS criticism of the correspondence theory of truth is directed primarily towards the metaphysical version of the theory. But this is not the only version that has been proposed and defended by philosophers. Perhaps a semantic version would not be subject to the criticisms of CBS? And what of the fact-based and object-based versions? Could it be that only one is incompatible with CBS, but not the other? Let us now finally turn to these issues, and see if a CBS-compatible correspondence theory is possible.

Part 4: Attempting a CBS-compatible correspondence theory of truth

In this section we will take a careful look at the CBS objections to specific versions of the correspondence theory of truth as outlined above. We will arrive at a version that does not seem to fall prey of these objections, and formulate such a version in CBS terms. I will conclude that this version is compatible with CBS, even though the CBS pragmatic theory of truth claims otherwise.

If you recall from the end of section 2.5, I listed five criteria for an analysis to be CBS-compatible:
1. That which is analyzed must be (treated as) a response/behavior
2. This response, as well as the surrounding stimuli and responses, must be understood (defined) in terms of the functions they have in the specific context
3. The analysis may only be done using CBS-terms (that would allow for prediction-and-influence with precision, scope and depth)
4. The stimuli and responses must be (at least in principle) observable (either publicly or privately), and cannot be hypothesized unobservable/abstract entities
5. Our analysis cannot be claimed to correspond to a pre-organized reality

We have seen that there are many different possible versions of the correspondence theory of truth. Some of the above criteria may provide objections only against specific truthbearers, truthmakers, or

[8] but of course only some concepts make sense when paired with certain concepts from the other categories.

relations, or specific combinations of these. I will therefore begin by examining each aspect separately. Upon finding that there are truthbearers, truthmakers and relations that are compatible with CBS I will then attempt to formulate a CBS-compatible correspondence theory of truth.

4.1 CBS criticism against truthbearers, relations and truthmakers of the correspondence theory of truth

4.1.1 Truthbearers

When it comes to truthbearers I believe the main objection from the perspective of CBS concerns whether they are observable or not (criterion 4). As has been laid out before, one of the principles of CBS is the idea that only that which can be observed (publicly or privately) makes sense to talk about. This philosophical standpoint rules out some of these truthbearers, for example that of beliefs. We cannot observe a belief directly. It is only ever inferred. What we can observe is a specific thought, in a specific instance. But from such an observation we cannot infer with confidence that that is a general belief that resides in our mind even when we are not thinking it. Propositions, if understood in an abstract sense, and not simply as written statements, are also not directly observable, and thus are ruled out. The issue is less clear on truthbearers such as thoughts, ideas and judgments. If these are considered as behaviors, and not a mental state, then they could in theory be acceptable. As we have seen before, thoughts can be understood using RFT. But if these terms are used to mean more of an abstract idea or thought, as in “the idea of snow being white”, rather than the specific behavior “John had the idea/thought that snow is white”, then they would be problematic. So the objection from CBS will depend on the usage of the terms. However, terms such as statements, utterances, and sentences are less problematic. Statements and utterances much more clearly refer to a specific behavior (or product of behavior), and sentences are stimuli that are even studied in relational frame theory. Thus, when it comes to the question of truthbearers, it seems like the principle of observability limits the list of CBS-compatible concepts to statements (and utterances, sentences) and, if understood as behaviors, thoughts (and ideas, judgments).

4.1.2 Truthmakers

When it comes to truthmakers I believe the main objection from the CBS perspective will be that of ontology (criterion 4 and 5). We can begin with terms like “fact”, “state of affairs”, and “condition”. One interpretation is that these terms refer to some abstract kind of entities. That they have some existence in some platonic world of ideas, where “facts” reside. The a-ontological stance would conflict with such an interpretation. Also, we cannot directly observe (publicly or privately) such abstractions. But another interpretation of these terms is that they simply refer to the observed (or to be observed) events that would confirm the truthbearer. Such an interpretation would be compatible with the CBS view.

What about the terms “event” and “object”? This depends on what we mean when we use these terms. If we make an ontological claim that objects and events are real in the sense of matching the pre-organized reality, then we are violating the a-ontological stance of CBS. But is such a claim necessary for the correspondence theory of truth? It is true that the theory is often expressed in general terms of correspondence with “reality”, which might suggest that terms like “object” and “event” should be understood in an ontological sense. But as we talked about in section 2.5.3, such an interpretation isn’t necessarily intended when using words like “reality”. Thus, the concepts “events” and “objects” may actually be compatible with CBS, as long as no ontological commitments are made (in the CBS sense). Another concept that may be compatible with CBS is “property”. In relational frame theory the term “formal properties” are used to describe physical features of a stimulus. If we

still fear that “property” risks being understood as an abstract concept we can, as we have seen previously, instead treat it as equivalent to the term “stimulus function”. Understood this way, the term “property” is compatible with CBS. Thus, when it comes to the question of truthmakers, it seems like all the terms are CBS-compatible, so long as they are understood as not making ontological claims.

4.1.3 Relations

Relations are a little more tricky to evaluate. The main reason for this is that the nature of a relation is rarely defined, and often only in metaphorical terms when done so, which is likely a consequence of the fact that relations by their very nature are quite abstract (David, 2016). One thing that CBS will object to in general, however, is if relations are postulated as some kind of abstract entities, i.e. the reification of these abstract relations (criterion 4). This would turn them into non-observable entities, and conflict with the a-ontological stance as well as the principles of observability. This does not seem to be the case for any of the proposed relations though.

But there are other ways that a relation may violate the a-ontological stance. For example, the relation of “copying” (as in a mental idea copies the external world) implies that a mental idea copies something of an *external reality*, which entails a dualistic view of the mental vs reality, implying something as being more real than something else. There may perhaps be some sophisticated way of claiming that the relation of “copying” could be interpreted in a way that does not entail such things, but for the purposes of this essay I will conclude that such a relation is not compatible with CBS. “Representation” is a similar relation. Often it may be used in the sense of a mental realm copying an external, objective realm. But it may be possible for this term to also apply to sentences representing something, i.e. in the sense of referring. So what about the relation of “referring”? As I mentioned in section 2.6.7, the term “refer” is not generally used in CBS, but can be naturally interpreted in CBS-terms as I outlined before. Another, closely related term is that of “signification”. To signify is generally understood as an umbrella term for all the different ways that one thing may stand in the place of another, among which is the relation of symbolizing. This is what relational frame theory is all about. Language understood broadly is synonymous with symbolizing, such that linguistic stimuli are sometimes called symbolic stimuli. As I see it, the relational frame of coordination (or more specifically the relation of stimulus equivalence) is exactly this kind of relation. The stimulus “written word ‘cat’” can signify the stimulus “visual object cat” if these stimuli are in stimulus equivalence (or frame of coordination) and the necessary contextual factors are present. There are other possible relations than those I have evaluated here, but this will suffice for our purposes. I conclude that quite a few different relations are compatible with CBS, and signification (and its synonyms) in particular fits well with RFT.

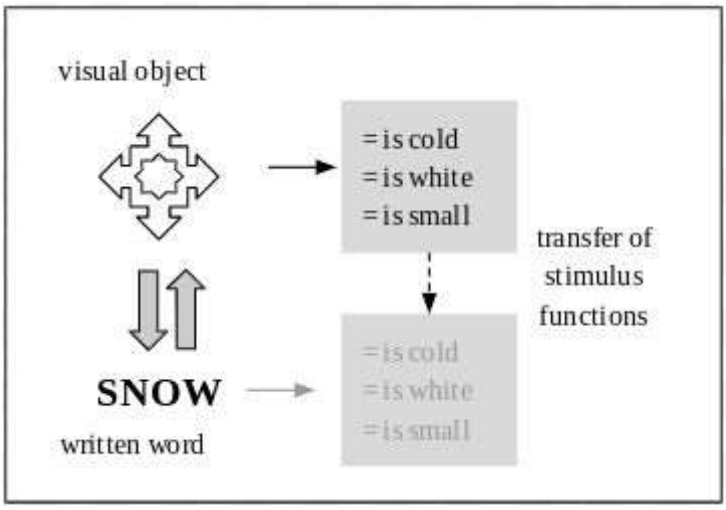
4.2 Attempting a correspondence theory of truth in CBS terms

So it seems like there are specific components that on their own are not in contradiction with CBS. But is it possible to combine them in a way that still makes sense and agrees with all criteria? Given what we have concluded above, a semantic object-based version of the correspondence theory seems most plausible, using concepts that on their own provide no direct contradiction to CBS. We can take a simple example and construct a version of the correspondence theory that involves the truthbearer “statement”, the relation “signifies”, and the truthmaker “object”. Would this give us a theory that says that “truth is when a sentence signifies an object”?

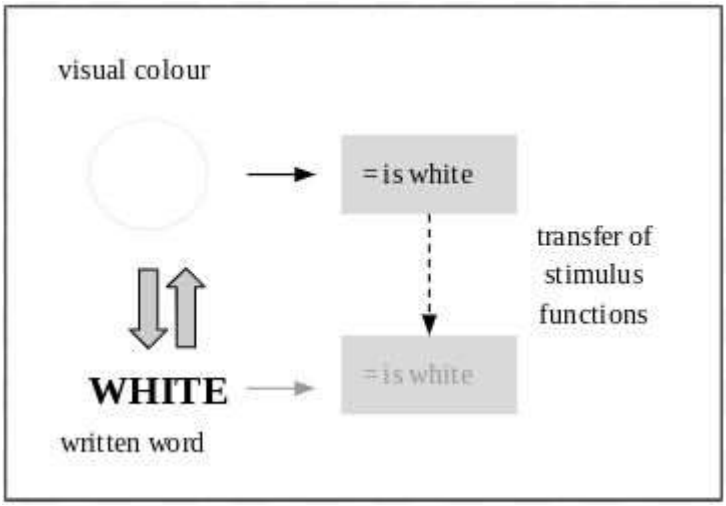
Well, if we think of what the object-based version of the correspondence theory was actually like, we may remember that it was a little more advanced than just saying that a statement signifies an object. It was that the *predicate* of a subject term corresponded to the *property* of the object to which the subject referred. The example I used previously was that of "Snow is white". In a semantic object-based version, the statement "Snow is white" is thus true if the object that "snow" signifies (or refers to) has the property specified by the predicate "is white". Let's see what happens when we try to express this version of the correspondence theory in terms of RFT, keeping in mind the five criteria presented earlier.

Beginning with the first criterion, a statement like "Snow is white" must be treated as a response to be CBS-compatible. What does this mean? Well it means we cannot think of this statement as some general idea that is not bound to any context, but must be understood as an utterance of a person who has a specific goal with expressing the statement and that this behavior is placed within the current and historical context which affects the functions of the statement and its parts. The statement will mean different things depending on the context. In the case of meteorologists discussing how the colour of the snow falling from the sky affects the reflection of light "Snow is white" will mean something different than in the case of a military operation where "Snow is white" is code for "The coast is clear". There are other contextual factors too. I mentioned before that we often leave out details when we communicate, and rely on the listeners learning history for a proper understanding. So when expressing "Snow is white", this could be meant as "Snow is always white", "Snow is often white", "Snow is sometimes white", or even "Snow is always white, unless coloured by some other substance". The correct interpretation will depend on context. When evaluating the truth of the statement "Snow is white", it is necessary for the precise meaning to be specified. In what follows I will therefore take the statement to be understood in the sense of "Snow is always white, unless coloured by some other substance." Let's now take a closer look at this statement.

"Snow is white" consists of a set of linguistic stimuli: "snow", "is" and "white". As a result of a learning history, the written stimulus "snow" is in stimulus equivalence with the visual object stimulus "snow", and the variation of such instances of snow that are encompassed by this stimulus equivalence is determined by the degree of generalization or discrimination. If saying the word "snow" has been reinforced for many different variations of the object "snow", then the word "snow" will be highly generalized for me, such that it includes a broad set of types of snow. The object snow has formal properties, such as being cold, being white, being small (relative to our body). These formal properties are stimulus functions of the stimulus "visual object snow", and because of the stimulus equivalence, these stimulus functions are transferred to the written stimulus "snow", such that this stimulus too has these stimulus functions.

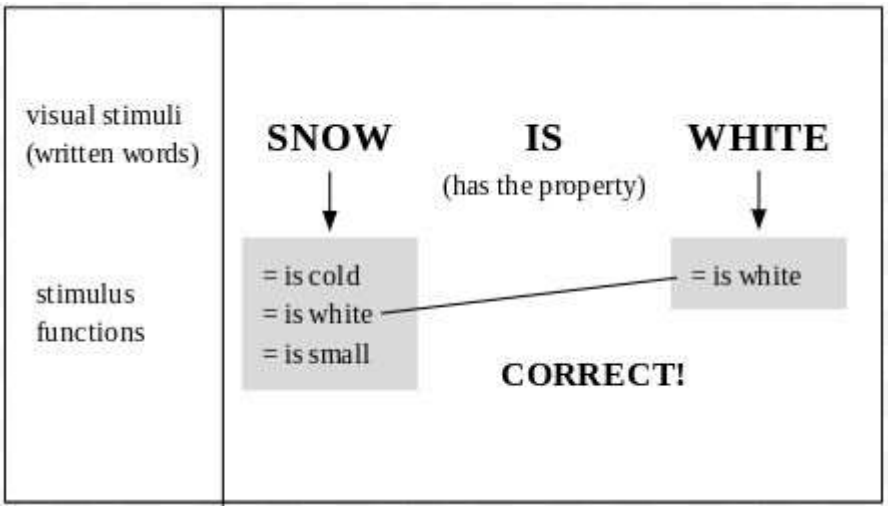


So far we have a CBS-description of the subject term in the statement “Snow is white”. Now let’s look at the predicate term “is white”. The predicate term consists of two linguistic stimuli “is” and “white”. If we begin with the term “white”, this written stimulus is in stimulus equivalence with the stimulus “the visual colour white”, again as a result of the learning history of the organism. The visual colour white has the formal properties of being white, which is thus a stimulus function of this stimulus. This stimulus function has been transferred to the written stimulus “white” because of their relation of stimulus equivalence.



The term “is”, on the other hand, is a relational contextual cue (Crel), which specifies which kind of relation is said to exist between two stimuli. It is a word that may have multiple meanings depending on context. Here we have contextual cues that imply how we are to understand “is”. Since what follows is the name of a colour, the word “is” is supposed to be understood as “has the colour”, or if we want we could describe it more generally as that of “having a property”, or “having the colour property”, to be more precise.[9] As I have said before, in CBS terms, “property” can be considered a synonym to “stimulus function”. Thus, the claim in “Snow is white” is that the stimulus “written word ‘snow’” has a stimulus function that matches the stimulus function of the stimulus “written word ‘white’”, in terms of colour. Which, as outlined, it does.

[9] or if we want, we could align this with the more general category of “frame of hierarchy” or “frame of coordination” depending on how these categories are defined.



This is of course just one example so far. But, according to the semantic object-based version of the correspondence theory of truth that we have now expressed in CBS-terms, the statement "Snow is white" is true. Expressing this theory in more general terms we arrive at a behaviorist version of the correspondence theory of truth, which I will call BCT for short.

> **Behaviorist correspondence theory of truth (BCT):** A statement is true if and only if the stimulus functions of the relata involved (i.e. the linguistic stimuli being related) match in the way specified by the relation (i.e. relational frame) between these relata.

From this follows that if the stimulus functions do not match, the statement is false. For example, in the statement "Snow is green", the stimulus functions of "snow" and "green" do not match. Thus, "Snow is green" is false.

The fact that we exclusively use CBS-terms (such as stimulus, stimulus function, stimulus equivalence or relational frame) means that we are fulfilling criterion 3. As mentioned before we are treating the statement as a behavior (criterion 1) and also understanding all stimuli involved in terms of the function they have in the specified context (criterion 2). We are not treating any of the stimuli as abstract entities, they are all observable (criterion 4), and we are not implying that the statement corresponds with a pre-organized reality and that it would therefore have a metaphysical privileged position in comparison to other possible analyses of the situation (criterion 5). This means that I believe that the correspondence theory, at least applied to the example of "Snow is white", is CBS-compatible.

I believe there may still be an objection from CBS on the more general level of whether this kind of truth is actually useful. The whole idea of the pragmatic theory of truth is that of usefulness (or successful working, or workability). It is often argued that in a clinical setting it is not useful to evaluate the correspondence truth (or "ontological truth") of the client's thoughts about herself or the world. That it is more important for the client to work towards her goals or values. But that cannot count as an argument against correspondence truth to ever be useful. For instance, it might be quite useful that a description of the client's problem corresponds to her actual problems, so that one can find an effective way of helping the client with her problems. I believe that correspondence truth is always *at least* a subgoal of a behaviorist analysis. Any behaviorist analysis involves the usage of words. If these words do not correspond to the stimuli involved in the specific situation, then our

analysis will in most cases be useless. Accepting the value of correspondence truth is thus not an abandonment of behaviorist principles. It is in service of them.

4.3 Expanding the application of the BCT

So far we have only analyzed a specific application of the behaviorist correspondence theory of truth (or BCT), namely to the statement "Snow is white". It seems likely that BCT will apply to any statement that has the form "stimulus 1" - relation - "stimulus 2" and where the stimuli and relation are simple, such that they don't seem to invoke any ontological claims. For example, "The food is warm", "Anna has the keys". But for the BCT to be considered an acceptable correspondence theory of truth it must apply to other types of statements as well.

We can begin with statements that do not have the same explicit structure but still involve simple stimuli and relations. One such example is "John eats", which may at first glance seem not to be an example of two relata being related via a relation. We talked about this example in section 2.6, and I explained how such a statement may be translated into "John is in the act of eating", or if we wish: "John has the property of eating", so that we can see the implicit structure more clearly. I believe this can be applied to any statement with a simple subject - predicate form. Thus, the BCT will work also for such statements.

What about more complex statements like "Placing the sweets in a box makes them hidden from sight"? We have to relata: "placing the sweets in a box" and "hidden from view", with the relation "makes". The second relatum "hidden from view" has pretty straightforward stimulus functions, but the first relatum is a little different. It is not just a simple stimulus, but an instruction that involves the combination of two subrelata: "the sweets" and "a box", and a (sub)relation between these: "placing in". Thus, we have to first apply this subrelation to the subrelata and investigate what the result would be if we followed the instructions. What we get by placing sweets in a box is simply "sweets in a box". This changes the original statement to "Sweets in a box are hidden from view". We can then look at how the stimulus functions of these simpler relata compare, and thus be able to conclude whether the sentence is true according to BCT in ordinary fashion.

But what about mathematical statements like "2 + 2 = 4"? If we begin untangling the structure here, we have two relata: "2+2" and "4" that are said to be in a relation of "=". Thus, the stimulus function of "2+2" must match that of "4" according to the relation of "=", which we may translate as "having the same quantity". Applying BCT, the first question is whether numbers or quantities are compatible with CBS, or whether they may be considered hypothesized unobservable entities. There is behaviorist research that has shown that not only humans, but also animals like rats can be reinforced to respond to quantities or numbers (e.g. Roberts, 2016), for example to always pick the third object in a row of objects. Thus, even if numbers have some level of abstraction compared to simpler perceptions like colour or shape, they are not hypothesized unobservable entities.[10] The second question concerns the relatum "2+2". This relatum contains an operation to be performed, namely that of adding 2 to 2. Is that really acceptable? We can compare this to the statement of placing sweets in a box, which we deemed acceptable despite an operation/instruction needing to be performed. So if both

[10] From cognitive research we know that the reason for the somewhat abstract nature of numbers is that they depend on core cognition systems (Carey, 2009) which takes simpler perceptual features as inputs. But such cognitive explanations aren't considered necessary or even useful to behaviorism. All that matters for something to be considered a valid stimulus is that we can indeed observe or respond to it.

numbers, and performing operations are acceptable, then the relatum "2+2" should likely be considered acceptable. It is simply a matter of applying our learning history of what "2" and "+" (or "is added to") mean. We can therefore find that "2+2" has the stimulus function of the quantity of four, which "4" has as well, thus the stimulus functions of the relata match according to the relation "has the same quantity".[11]

So far it seems that BCT may potentially apply to quite a broad set of statements, but there is one thing we need to address. The different statements have so far been evaluated by comparing the stimulus functions of the written stimuli, i.e. the words. For example, when evaluating "Snow is white", all we needed to do in the end was to say that the written word "snow" has stimulus functions that matched those of the written word "white" (according to the relation "has the colour"). Thus, the truth of the statement relied on a coherence between the stimuli involved. But the point of BCT was to be a *correspondence* theory of truth, not a *coherence* theory. It is therefore worth pointing out that the coherence between the written stimuli still relies on the fact that the stimuli correspond to (or are in stimulus equivalence with) with other stimuli that are under interest. The fact that the application of BCT so far has seemed like a coherence theory is actually because the stimulus functions involved are already known. No additional observation of stimuli referred to have been necessary. We shall therefore look at examples where such observations (or at least additional information) is necessary. But it is worth pointing out that a coherence theory of truth thus follows as a special case of BCT, which explains how such seemingly diverse types of truth really are the same thing at the core. It also points at an interesting way of comparing analytic and synthetic truths, the distinction originally made by Kant (1781 [1998]), since some of the above examples have involved mere analysis, while others have involved operations. But this question lies outside the scope of this essay.

Let us now therefore turn to cases that contain unknown stimulus functions. One such example could be the statement "The ball in this bag is blue". We have two relata: "the ball in this bag" and "blue", and the specified relation between them is "has the colour". The relatum "the ball in this bag" contains a number of words, where the word "this" is an indexical that is in stimulus equivalence with whatever is indicated in the specific context. Thus, we may assume that there is a bag in front of the speaker that is somehow indicated. Let's say that in the specific context, "the ball in this bag" is something we are only acquainted to through feeling around in the bag with our hands. We can tell that this stimulus has stimulus functions such as "is round", "is hard", "has a smooth surface", however we have not been able to make a visual observation of the ball, so we do not know the colour of the ball. Thus, we cannot assess the truth of the statement analytically. We need to make additional observations (or be told about observations that others have made of the ball) so that the stimulus function under interest is added to (or transferred to, if we learn through linguistic input) the stimulus. The statement will then be either true or false depending on the result of this observation. If it turns out to be blue, then the statement is true, if it is not blue, then the statement is false.

What about the truth of statements that involve more abstract entities? For example, "God exists". Is that something that can be made sense of by the BCT? Well, let's begin with looking at a statement like "Horses exist". The acceptability of such a statement will depend on what is implied by "exist" in this case. As outlined in the section on the a-ontological stance, "ontology" in terms of the "explicit specification of concepts to be used in a domain" is okay. So if we are simply stating that horses are a

[11] It is not clear, however, how this generalizes to cases with quantities/numbers like 1000, which cannot be perceived in the same manner as 1 up to perhaps 10.

type of stimulus that we can talk about in behaviorist analyses, then that is fine. Another interpretation is that "exist" has the meaning "has physical form", which then becomes a question of comparing the stimuli "horse" and "physical form", in terms of the relation "has". That might also be okay. But other interpretations, that put ontological commitments into the word "exist" would be problematic. When talking about "God" instead of horses, we have an additional difficulty in finding a way of interpreting "God" as referring to something observable. Some may argue that "God" can be observed in terms of the beauty of the world, or a certain feeling in their body. If one thinks of "God" as being the love in the world, then it might also imply that "exists" should be interpreted as simply having those features. So, there doesn't seem to be any principled objection to statements of this kind from BCT, it simply depends on what functions these stimuli are proposed to have in specified contexts.

Let us now look at logical statements like "If p, then p or q". This involves yet another layer of complexity since p and q are statements in themselves, and are considered to have a truth value of their own. So we are essentially relating statements. For someone unfamiliar with these expressions, it is worth pointing out that what it is saying is that "If the statement 'p' is true, then the statement 'p' or the statement 'q' is true" (where 'or' in this context is supposed to be understood as and/or). Thus, we are setting up a context where we are saying that the statement 'p' is true (the first relatum), and asking whether this means that either the statement 'p' or the statement 'q' is true in this context (the second relatum). Another way of expressing the statement to make it explicitly follow the structure of stimulus - relation - stimulus, could be "'p' is true has the consequence that 'p or q' is true". Given that we know that the stimulus function concerning truth for the statement 'p' is 'true', then this means that the same stimulus functions apply to the statement 'p' in the expression 'p or q'. Given our knowledge (our previous learning history) of statements with the logical connective 'or', we can discern that the expression 'p or q' therefore also must have the stimulus function 'true'. Thus, the original statement is true according to BCT.

Another type of statement that may be of interest is paradoxical statements like "This statement is false". Here we have the relata "this statement" and "false". When applying BCT, we have to establish what the stimulus "this statement" is in stimulus equivalence with, so that we can figure out what stimulus functions this stimulus has. When we do so we find that "this statement" is in stimulus equivalence with the stimulus "this statement is false", but what are the stimulus functions of this statement? Well, that's what we wanted to find out to begin with, so this is unknown. What we get is an expanded statement "[This statement is false] is false", in order to understand this statement, we once again have to find the stimulus equivalence of the first relatum, which means we have to look at what "this statement" is in stimulus equivalence with in this case too. This will just lead to an ever expanding statement like "[[This statement is false] is false] is false". According to BCT, the self-referential nature of the statement means that we never find out the necessary stimulus functions to evaluate the statement. This could perhaps be an acceptable way of explaining the apparent paradox.

Part 5: Conclusion

I believe that the above shows that the correspondence theory of truth, when expressed in CBS-terms, is not in contradiction with the actual principles of the philosophy of functional contextualism. In fact, I believe that it is an important subgoal of the goal of CBS. Since the formulation of the pragmatic truth criterion is not coherently expressed or argued for as it stands, I believe that a better formulation would be to state that "successful working" is an important *goal*, not a criterion of truth. Further that

the correspondence theory of truth should be embraced by CBS, and seen as an important subgoal for any analysis.

Further, I believe that the above means that cognitive science cannot be criticized on the grounds of using the correspondence theory of truth. However, it is still possible that objections against its use of non-observable entities still hold. It might also be possible that while the use of the correspondence theory, per se, is legitimate and in accordance with CBS, the way that truth is sought in practice might be problematic. If cognitive research considers the establishing of true statements is *enough* for also influencing events, then this is mistaken. Whether this is the case or not is another question. It could be a mischaracterization of cognitive science as being interested in correspondence truth with no further purpose. There might also be a further goal behind this search for truth that might not be explicitly stated. Hence, the importance of making the philosophical assumptions and goals explicit, which is an amazing feat of CBS, and deserves credit regardless of possible flaws.

Although the main purpose of this essay was to investigate the correspondence theory from a CBS perspective, I believe that the theory characterized in CBS terms may be useful also outside of CBS. For one I believe it can solve the question of how truth is thought of as an asymmetrical relation, while "correspondence" is a symmetrical relation (David, 2016). As previously shown, relational frames only appear asymmetric because there is an implicit (historical) context that tells us to think of the relational frame in one direction and not another in certain (current) contexts, while the relation itself is still symmetrical. It might also be useful for investigating questions of analytic vs synthetic truths, or a priori or a posteriori truths.

The CBS-version of the correspondence theory of truth that has been outlined may still have issues other than those that have been discussed. Questions remain regarding how this version would apply to even more complex and diverse cases. It may also have to answer to objections posed to the correspondence theory more generally in philosophy. These things may be a topic for a later project.

References:

Biglan, A., & Hayes, S. C. (1996). Should the behavioral sciences become more pragmatic? The case for functional contextualism in research on human behavior. *Applied and Preventive Psychology: Current Scientific Perspectives*, 5, 47–57, http://dx.doi.org/10.1016/S0962-1849(96)80026-6.

Biglan, A., & Hayes, S. C. (2016). Functional contextualism and contextual behavioral science. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (p. 37–61). Wiley-Blackwell

Bourget, D. and Chalmers, D., 2014, 'What Do Philosophers Believe?', *Philosophical Studies*, 170: 465-500.

Carey, S. (2009). *The origin of concepts*. Oxford University Press

Catania, A. C. (2007). *Learning. Interims.* (4th edition). New York: Sloan Publishing

Chomsky, N. (1959). Verbal Behavior by B. F. Skinner [Review]. *Language*, 35, 26-58. 10.2307/411334

David, M. (2016). The Correspondence Theory of Truth. *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/fall2016/entries/truth-correspondence.

Hayes, S. C. (1993). Analytic goals and the varieties of scientific contextualism. In: S. C. Hayes, L. J. Hayes, H. W. Reese, & T. R. Sarbin (Eds.), *Varieties of scientific contextualism* (pp. 11–27). Oakland, CA: Context Press/New Harbinger.

Hayes, S. C., Barnes-Holmes, D., & Wilson, K.G. (2012). Contextual Behavioral Science: Creating a science more adequate to the challenge of the human condition, *Journal of Contextual Behavioral Science*, Volume 1, Issues 1–2, Pages 1-16, ISSN 2212-1447, https://doi.org/10.1016/j.jcbs.2012.09.004.

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations and a behavior analytic view of the purposes of science. *The Behavior Analyst*, 9, 175–190.

Hayes, S. C., Hayes, L. J., & Reese, H. W. (1988). Finding the philosophical core: A review of Stephen C. Pepper's World Hypotheses. *Journal of the Experimental Analysis of Behavior*, 50, 97–111, http://dx.doi.org/10.1901/jeab.1988.50-97.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy: The process and practice of mindful change* (2nd ed.). New York: Guilford Press.

Hirsch, E. (2011). "Introduction". *Quantifier Variance and Realism: Essays in Metaontology*. Oxford University Press. p. xii. ISBN 978-0199732111.

James, W., (1907) [1975]. *Pragmatism: A New Name for Some Old Ways of Thinking*, New York and London: Longmans, Green & Co.; reprinted Cambridge, MA: Harvard University Press, 1975.

Kant, I., (1781) [1998], *The Critique of Pure Reason*, P. Guyer and A.W. Wood (trans.), Cambridge: Cambridge University Press.

Levin, M. E., Twohig, M. P., & Smith, B. M. (2016). Contextual behavioral science: An overview. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (p. 17–36). Wiley-Blackwell.

Pepper, S. C. (1942). *World hypotheses: A study in evidence.* Berkeley: University of California Press.

Peirce, C. S., (1878) [1986]. "How to Make Our Ideas Clear", *Popular Science Monthly*, 12(January): 286–302; reprinted in *Writings of Charles S. Peirce: A Chronological Edition* (Volume 3: 1872–1878), C. Kloesel, M. Fisch, N. et al. (eds.), Bloomington, IN: Indiana University Press, pp. 257–276.

Roberts, W.A., (2016) Rats respond where it counts. *Learning and Behavior 44*, 101–102. https://doi.org/10.3758/s13420-016-0229-y

Russell, B. (1912). *Problems of Philosophy,* reprinted at Oxford: Oxford University Press 1971.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research*, 14(1), 5–13. https://doi.org/10.1044/jshr.1401.05

Skinner, B. F. (1938). *Behavior of organisms*. New York: Appelton-Century-Crofts.

Skinner, B. F. (1957). *Century psychology series. Verbal behavior*. Appleton-Century-Crofts. https://doi.org/10.1037/11256-000

Törneke, N. (2010). *Learning RFT. An introduction to relational frame theory and its clinical implication*. Context Press/New Harbinger Publications.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20(2), 158–177. https://doi.org/10.1037/h0074428

Zettle, R. D., Hayes, S. C., Barnes-Holmes, D., & Biglan, A., (Eds.), (2016) *The Wiley handbook of contextual behavioral science* (p. 17–36). Wiley-Blackwell.