

Verbal Reports in the Experimental Analysis of Behavior

Francisco Cabello Luque*

Universidad de Murcia, España

Denis O'Hora

National University of Ireland Galway, Ireland

ABSTRACT

This paper provides a review of the role of verbal reports for the study of human verbal behavior from a behavior-analytic perspective. Though typically associated with cognitive research, verbal reports have long been considered an important dependent variable within behavior analysis. Behavior analysts are particularly wary of being lured into equating an individual's report of the controlling variables in a context with the actual variables at work. Indeed, the complexity of verbal reports requires complex tools and procedures and, in the current manuscript, we describe two such tools (protocol analysis and the silent dog method), and review the current literature of studies using them. In general, these tools have demonstrable utility as tools for the analysis of verbal behavior and results obtained thus far are interesting, but there is a lack of standardization across procedures that hinders cumulative progress. The current review highlights the strengths of these tools in permitting a functional analysis of self-generated rule following and suggests future research to enable the development of standardized approaches to self reports in behavior analysis.

Key words: verbal reports, protocol analysis, silent dog method, verbal behavior.

Novelty and Significance

What is already known about the topic?

- Verbal reports are a common dependent variable in Cognitive Psychology.
- The validity of verbal reports, however, has been critiqued within behavior analysis.

What this paper adds?

- A historical perspective of the behavior-analytic interest regarding verbal reports.
- A description of protocol analysis and of the silent dog method, and a brief review of the studies that have used these tools.
- A discussion of the strengths of both tools for the analysis of verbal behavior.

When exposed to experimental tasks, human participants typically engage in verbal behavior such as counting, describing what they are doing or thinking out aloud, even without being instructed to do so. Also, after they complete a task, it is common to ask them to explain why they responded the way they did, or to justify their reasoning.

For a long time, psychologists have used this verbal behavior as a means to describe and even explain the participants' performances; for example recording what they were saying, asking after the experiment to describe what they had done, or applying a post-experimental questionnaire. Curiously, some of the authors who established the basis for behavior analysis used such practices: Watson is recognized as one of the first psychologists to use "think aloud" techniques (Watson, 1920) and Skinner makes reference to this topic in several of his books (Skinner, 1957, 1984), when discussing

* Correspondence concerning this article: Francisco Cabello Luque, Departamento de Psicología Evolutiva y de la Educación, Facultad de Psicología, Universidad de Murcia, Campus de Espinardo, 30100, Murcia, España. Email: fcabello@um.es

verbal behavior under the control of private events and the difficulties to establish a referent for what participants are saying during experiments.

However, the use of verbal reports in the experimental analysis of behavior was very uncommon for a variety of reasons, such as the interest in differentiating from mentalism, or the hypothesis that the principles of human behavior should be studied using non human animals such as rats, pigeons and monkeys. In this context, what participants said during of after the experiment was disregarded as non-valid data. It was not until the decade of the 90s, and its renewed interest on the experimentation with human participants, that behavior analysis fully considered the value of verbal reports. Since then, new theoretical and methodological frames have been proposed to use these reports as valid data for the analysis of human behavior.

In this paper, we will review the role of verbal reports from a behavioral perspective, starting with a brief historical review that will show both the empirical evidence obtained using verbal reports and the criticisms that were legitimately raised against their use as valid data. Then we will present recent proposals that attempt to overcome these critiques through the rigorous collection of verbal reports under conditions of minimum listener control. Lastly, we will discuss the current state of the topic and the problems that still need to be solved in order to use what participants say, express and manifest in our laboratories as part of our research data.

THE USE OF VERBAL REPORTS AS DATA: A HISTORICAL PERSPECTIVE

Verbal reports in Psychology

Verbal reports have been applied to numerous topics in Psychology, such as (without being exhaustive) the formation of beliefs and attitudes, their relation to behavior, problem-solving strategies, the application of analogical reasoning, differences between experts and novices, academic problem-solving, strategies for decision-taking, thinking processes, creative thinking, second language learning, comprehension processes, logical reasoning, metaphorical thinking, perspective taking, theory of mind development, executive function, self-control, the diagnostic process of doctors, education, the relation between cognitive processes and academic achievement, or even to clinical areas such as phobia or anxiety (see Crutcher, 1994; Ericsson & Simon, 1980, 1993; or more recently, Berk, 1999; Conrad, Blair, & Tracy, 1999; Ericsson, 2003; Winsler, Fernyhough, & Montero, 2009).

Behavior analysts have traditionally been particularly skeptical of verbal reports, since they are often construed as evidence of the status of cognitive processes that mediate the interaction between the organism and the environment. Even though behavior analysts accept that verbal reports constitute a performance that is under the influence of a person's behavioral history and the current context, they are very wary of any 'slippery slope' that would lead them to equate verbal reports with the operation of internal mechanisms.

In attempting to map out appropriate uses of verbal reports within behavior analysis, Perone (1988) described how verbal reports can be used by behavior analysts

for the following goals: (1) To provide the only practical means of observing certain forms behaviors (e.g., to measure the consumption of illicit drugs through questionnaires); (2) To collect data on characteristics of behavior that are defined in terms of verbal responses to a set of questions, such as the measurement of attitudes and beliefs; (3) To directly evaluate the content of the verbal reports themselves, understanding that they play a role in behavior and that their content is relevant; (4) To produce explanations based on the content of verbal reports, based on the assumption that they have a causal role on behavior.

Verbal reports in the Analysis of Behavior

As mentioned above, verbal reports were used by early behaviorists (Watson, 1920). However, they were disregarded for theoretical and methodological reasons. For example, Critchfield and Epting (1998) and Critchfield, Tucker, and Vuchinich (1998) note that verbal reports were understood to indicate mentalism, which was something behavior analysis should reject. Besides such reports lacked practical utility when analysing the behaviour of nonhuman animals, the typical participants in behaviour analysis; for example, Buskist and Miller (1982) report that 93% of papers published in the *Journal of the Experimental Analysis of Behavior* (JEAB) by the end of the 70s included non-human participants, a very different situation from general Psychology at the same time, in which studies with humans were 10 times more common than those with animals (Miller, 1985).

These and other reasons led to a situation in which verbal reports were left out of the most common experimental arrangements, were not collected or interpreted systematically, were not frequently used, and were not considered particularly useful. The following quote, taken from Perone (1988), is very clear:

I first became aware of this when Alan Baron submitted a paper to the JEAB. The work was concerned with conditioned reinforcement of human observing responses [...]. One of the reviewers commented as follows: "I was surprised that no verbal reports -summarizing what the subjects said they were doing and why- were presented. If available, they should be discussed". This struck me as an odd request. Why ask the subjects what they did? (p. 71).

This situation, however, changed gradually by the late 80s and early 90s, when a renewed interest for human research began in behavior analysis (Hyten & Reilly, 1992; Navarick, Bernstein & Fantino, 1990; Perone, 1988). Among the reasons for this "renaissance" of human experimentation, the early research on equivalence relations, the literature on insensitivity to contingencies, and the work on rule-governed behavior have been cited (Hayes, 1989; Luciano, 1992, 1993). These disparate areas shared a fundamental idea: that verbal behavior could change the way in which human organisms interacted with the environment, thus challenging the continuity hypothesis and proposing that non-human animals were not an adequate model for certain areas of research (Bentall, Lowe, & Beasty, 1985; Dymond, Roche, & Barnes-Holmes, 2003; Dugdale, 1988; Hayes, 1987; Horne & Lowe, 1996; Lowe, Harzem, & Huges, 1978).

As a consequence of the interest on verbal behavior, what participants said became again a relevant dependent variable that could produce useful data, and with increasing frequency, empirical studies reported the use of verbal self-reports. By the end of the 90s and in a few years, for example, Catania and Shimoff (1998) considered these reports an important issue for the analysis of verbal behavior; Fukui (2002) included self-reports in a list of "recent topics" for the analysis of human behavior, and as we will see, entire issues of journals were devoted to examine how they could be used in the context of behavior analysis (Schlinger, 1998).

THE TIME AND PLACE FOR VERBAL REPORTS

Even though behavior analysts warmed to the prospect of analyzing verbal reports as experimental data during the 90s, it soon became clear that the go-to approaches to collecting such reports suffered from methodological limitations (see Critchfield, Tucker, & Vuchinich, 1998; Hayes, 1986; Luciano, 1993; Perone, 1988; Shimoff, 1986; but also Ericsson & Simon, 1980, 1993). The most common methods employed to collect verbal reports at this time were to obtain post-experimental reports or to prompt verbal reports while an experiment was in progress.

In the first group of studies, post-experimental reports were obtained from participants who had completed an experiment by asking them to describe any rules they have been following, under the assumption that these reports will somehow correspond to the private verbal behavior during the experiment. For example, Dube, Green, and Serna (1993) demonstrated the emergence of equivalence relations using auditory stimuli, specifically musical tones generated by a computer. Once baseline relations were trained and equivalence was tested, they asked participants to report any verbal rule they had been following during the experiment. All participants demonstrated equivalence and there was a consistent relation between verbal reports and task performance, with the behavior of participants during the experiment closely matching the post-experimental reports.

The following criticisms have been raised regarding this type of procedure for the collection of verbal reports: (1) Verbal reports can simply be under the control of the questions made by the experimenter; (2) Public verbal reports could be not reliably related to private verbal behavior during the task; (3) It might be difficult to conclude anything about the verbal behavior of participants during the whole experiment using just a single measure at the end of the task; (4) Verbal behavior during the experiment might not be related to task performance, and thus becomes irrelevant.

A second group of studies are those in which verbal reports are prompted several times during the task, and not simply at the very end. For example, Moreno, Ribes, and Martínez (1994) conducted an experiment in which second-order discriminations were trained using geometrical figures as stimuli, so that participants had to establish responding according to relations of sameness (choosing the comparison with the same color and shape than the sample), difference (choosing the comparison with the same color but different shape) and opposition (choosing the comparison that is different in color and shape). During some trials, participants were asked to describe their strategies by completing a sentence like "when at the top there is a _____ and in the center there

is a _____, I choose _____” using several response options. Participants’ verbal descriptions matched their performances on the vast majority of trials.

The prompted report approach had the following limitations: (1) Obtaining verbal reports can interrupt task development and thus affect task performance; (2) There is a possibility that asking participants to report what they are doing introduces changes in their performances compared to other participants who are not requested to do so; (3) It could be that verbal behavior is irrelevant for task performance.

The foregoing criticisms, however, do not mean that verbal reports do not have a place in the analysis of behavior. On the contrary, there is a clear consensus about the fact that these measures could be useful to analyze and investigate human thinking and cognition, to allow for an analysis of the antecedents and consequences of verbal behavior, or to make public events that happen privately. What these criticisms mean is that there are some problems that need to be addressed, and the use of verbal reports should not be done solely on the basis of conceptual and theoretical elaborations, but through a process of empirical validation that maximises their validity (Critchfield, Tucker, & Vuchinich, 1998). In other words, the question is not whether verbal reports can be used or not, but how they should be used. In the last two decades, two tools have been developed that address the criticisms raised above and that might allow for (a) obtaining verbal reports in adequate conditions, (b) determining their validity, and (c) analyzing them adequately. These tools are the protocol analysis and the silent dog methodological controls, and are described below.

PROTOCOL ANALYSIS

What is protocol analysis?

The first significant advance for the use of verbal reports in behavior analysis comes from cognitive psychology. Although “think aloud” procedures were used from the beginning of the 20th century (Bulbrook, 1932; Duncker, 1945; Watson, 1920), it was not until the work of Ericsson and Simon (1980, 1993) that these procedures became standardized and a coherent theoretical framework was proposed, called protocol analysis. The protocol analysis method was developed, in part, to deal with the limitations of post-experimental reports and prompted reports, which had also been identified in the mainstream psychological literature. In particular, Ericsson and Simon sought to reduce audience effects on verbal reports.

During protocol analysis, subjects are asked to “think aloud” and to verbalize their thoughts at the same time they are engaged in the experiment as if they were alone, and after that verbal reports are systematically transcribed, divided into segments and analyzed. The following is a typical example of the method, based on Ericsson and Simon (1993):

1. Before starting the experiment, participants receive instructions about the experimental task, and are also given specific instructions about the “thinking aloud” requirement: “In this experiment we are interested in what you think about when you find answers to some questions. In order to do this I am going to ask you to think aloud as you

- work on the problem given. What I mean by think aloud is that I want you to tell me everything. I would like you to talk aloud constantly, I don't want you to try to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself."
2. After receiving instructions, subjects also receive practice at "thinking aloud" using simple exercises such as mental calculations. These exercises are used to prompt the participants to think aloud and to shape their verbal behavior, as "So that you understand what I mean by think aloud, let me give you an example. Assume I asked you 'How much is 127 plus 35?'. Now think aloud so I can hear how you solve this problem."
 3. Once that the training is completed, subjects are exposed to the experimental task, while prompted to keep "thinking aloud" if they remain silent for some time. The participants' verbalizations are recorded during the experiment (e.g. using audio tapes).
 4. The following step involves the transcription into text of these verbalizations. Once that the content of the tapes is converted to written form, it is divided into several segments (e.g. speech sentences, or trial by trial).
 5. These written segments are then assigned to different categories devised by the experimenter, usually by at least two independent raters. The categorizations of the different raters should be checked for inter-rater reliability.
 6. Then, the categorized protocols are ready to be analyzed. In the case of cognitive psychology, the protocols are used to build an information-processing model that is later simulated in the computer (Newell & Simon, 1972).

The key issue in protocol analysis, and what makes it specially interesting, is that it goes beyond instructing participants to "think aloud" and attempts to fully systematize the process to obtain verbal reports from experimental participants under controlled conditions of minimal audience, and the analysis of such reports (Austin & Delaney, 1998).

The Behavior Analysis interest in protocol analysis

Hayes (1986) was the first to review the protocol analysis method and to propose that it could be useful for the study of rule-governed behavior and self-rules. This idea was later discussed by Austin and Delaney (1998), Critchfield and Epting (1998), and Hayes, White and, Bissett (1998) in a special issue of *Analysis of Verbal Behavior*, and latter commented by Potter (1999) and Normand (2001). In all cases, the authors agree that protocol analysis could be of interest for the analysis and investigation of human thinking and of private verbal behavior. More importantly, this interest soon translated into a series of empirical studies that, although related to different research topics and using different procedures, have applied the protocol analysis method within the framework of behavior analysis. Table 1 is a listing of the main characteristics of these studies and of their main results.

An inspection of these studies shows an important variability in aspects that might be of importance (see Normand, 2001). First, experimental procedures related to the "think aloud" requirement are different: some studies closely follow Ericsson and Simon's recommendations indicated above, that consist of specific instructions to

Table 1. Characteristics of published studies that have used protocol analysis from a behavioral perspective. For each study, it is indicated what it was aimed at, the type of participants, how verbal reports were obtained, how verbal reports were analyzed into categories, and the main results.

	Aimed at	Determining the implication of verbal behavior for the emergence of equivalence classes
	Participants	Adults (students)
Wulfert, Greenway, & Dougher (1991)	Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter remained in the room for 3-5 minutes
	Coding categories	<ul style="list-style-type: none"> • Relational responding • Common physical features • Stimulus compounds • Other verbal utterances
	Main results	Found significant relations between verbal reports and task performance; participants who did not demonstrate equivalence showed more common physical features and stimuli compounds, while those demonstrating equivalence produced more relational responses
	Aimed at	Determining the implication of verbal behavior for the emergence of third-order equivalence classes
	Participants	Adults (students)
Wulfert, Greenway, & Dougher (1994)	Procedure for verbal reports	Specific instructions to think aloud, practice exercises. Experiment was not present in the room
	Coding categories	Varied for each experimental phase, but similar to the latter study
	Main results	Found significant relations between verbal reports and task performance; researchers were able to better determine how training was functioning and learned about extraneous sources of control
	Aimed at	Examining mediational verbal behavior in the distinction between selection-based and topography-based behavior
	Participants	Adults (students)
	Procedure for verbal reports	Specific instructions to think aloud, practice exercises. Experimenter was present in the back of the room
Potter, Huber, & Michael (1997)	Coding categories	<ul style="list-style-type: none"> • Tacts • Repeated facts • Repeated intraverbals • New intraverbals • Repeated test tact • Repeated test intraverbal • Other verbal utterances
	Main results	In general, repeated facts and repeated intraverbals were related to correct trials. On the contrary, new intraverbals and other verbal utterances were related to failed trials
	Aimed at	Examining the blocking effect in a stimulus equivalence procedure
	Participants	Adults (students)
Rehfeld, Dixon, Hayes, & Steele (1998)	Procedure for verbal reports	Specific instructions to think aloud. Experimenter was not present in the room
	Coding categories	<ul style="list-style-type: none"> • Relations between sample and comparison stimuli • Descriptions of comparison stimuli • Variables unrelated to the experiment • Other verbal utterances
	Main results	Participants who emitted more relations between sample and comparison stimuli during tests, obtained worse performances and a weaker blocking effect. Participants who emitted descriptions or other utterances obtained better performances. This relation was not consistent

Table 1. Characteristics of published studies that have used protocol analysis from a behavioral perspective. For each study, it is indicated what it was aimed at, the type of participants, how verbal reports were obtained, how verbal reports were analyzed into categories, and the main results (cont.).

	Aimed at Analyzing the influence of instructional history on rule-following
	Participants Adults (students)
	Procedure for verbal reports Specific instructions to think aloud. Experiment was not present in the room
	Coding categories <ul style="list-style-type: none"> • Reinforcement • Performance • Correct contingencies • Incorrect contingencies • Aversive/fatigue • Reading from screen • Experimenter directed • Unrelated to task
	Main results Participants who emitted utterances about reinforcement or correct contingencies needed less trials to complete the task
Dixon & Hayes (1998)	Aimed at Relating verbal behavior to the generalization and long-term retention of equivalence classes
	Participants Adults (students)
	Procedure for verbal reports Specific instructions to think aloud. Experimenter was not present in the room
	Coding categories <ul style="list-style-type: none"> • Relations between sample and comparison stimuli • Description of physical properties of individual stimuli • Reinforcement or its removal • Variables not related to the experiment • Variables irrelevant to the stimuli arrangement • Silence
	Main results Relating stimuli and descriptions were the most common categories. Most participants showing generalization had relating stimuli as the most frequent category. During retention tests, relating stimuli is related to better performances. The relation, however, is not consistent
	Aimed at Studying the use of median lines as a method to increase the validity of data analysis
	Participants Adults (students)
	Procedure for verbal reports Specific instructions to think aloud, practice exercises, reminders to think out loud. Experimenter remained in the room
	Coding categories Not specified
	Main results Verbal reports suggested that median lines were not related to the validity of the tendencies that participants inferred from statistical data
Rehfeldt & Hayes (2000)	Aimed at Relating stimuli and descriptions were the most common categories. Most participants showing generalization had relating stimuli as the most frequent category. During retention tests, relating stimuli is related to better performances. The relation, however, is not consistent
	Participants Adults (students)
	Procedure for verbal reports Specific instructions to think aloud, practice exercises, reminders to think out loud. Experimenter remained in the room
	Coding categories Not specified
	Main results Verbal reports suggested that median lines were not related to the validity of the tendencies that participants inferred from statistical data
Normand & Bailey (2006)	Aimed at Studying the use of median lines as a method to increase the validity of data analysis
	Participants Adults (students)
	Procedure for verbal reports Specific instructions to think aloud, practice exercises, reminders to think out loud. Experimenter remained in the room
	Coding categories Not specified
	Main results Verbal reports suggested that median lines were not related to the validity of the tendencies that participants inferred from statistical data

Table 1. Characteristics of published studies that have used protocol analysis from a behavioral perspective. For each study, it is indicated what it was aimed at, the type of participants, how verbal reports were obtained, how verbal reports were analyzed into categories, and the main results. (cont.)

Aimed at	Investigating the role of common names for the emergence of equivalence classes
Participants	Adults (students)
Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter was not present in the room
Coding categories (Garcia & Rebfieldt (2008))	<ul style="list-style-type: none"> • Description of matching between stimuli on the basis of names or numerical values • Single names or numerical values but without mention of matching • Description of matching but without names or values • Matching on the basis of physical features • Matching on the basis of contingencies • Irrelevant utterances • Silence
Main results	Participants were taught, before being exposed to the equivalence task, to give class-consistent names to the stimuli or to emit common fixed ratio responses in the presence of stimuli. When instructed to think aloud, verbal reports demonstrated that this pre-task training was successful and that assigning a common name or a common FR response made the experimental stimuli functionally equivalent prior to any training
Aimed at	Comparing concurrent and retrospective self-reports about a serial task
Participants	Adults (students)
Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. A control condition was used that completed the task in silence. Experimenter was not present in the room
Coding categories (Plancarte, Moreno, Hickman, Arroyo, & Cepeda (2013))	<ul style="list-style-type: none"> • Evaluation utterances • Counting • Comparing and/or copying • Descriptions • Emotional utterances • Attentional • Other
Main results	Differential effects were found based on the use of concurrent or retrospective self-reports. Verbal reports facilitated task performance in general, but the type of verbal reports were significantly different if they were obtained concurrently or retrospectively, and their relation to task-solving strategies was also different
Aimed at	Testing whether Access to B stimuli were necessary to pass CA equivalence test, given that AB and BC relations were trained
Participants	Adults (students)
Procedure for verbal reports	Instructions to read aloud the stimuli before the main task. Experimenter was not present in the room
Coding categories (Dickins (2015))	<ul style="list-style-type: none"> • Utterances containing A stimuli • Utterances containing B stimuli • Utterances containing C stimuli
Main results	Participants who emitted utterances containing B stimuli during CA tests obtained better performances. This pattern was also obtained when varying the type of test and the training sequence. However, these results were not found for all participants

think aloud, practice exercises, and reminders to think aloud, but others do not and use different instructions or omit warm-up exercises before the experimental task. Second, instructions provided to participants are widely different among studies, with some experiments using very detailed instructions about what participants should do, and others using minimal instructions in order not to shape possible verbal behavior. Third, in some of the studies the experimenter is present in the same room as the participant, but, in others the experimenter is absent, constituting different audience conditions across studies. And fourth, there are no standard categories for the analysis of verbal reports, and each study uses its own scheme. Moreover, the impact of these different procedures is typically not discussed, and thus similar inferences are drawn from procedures that differ in the way that verbal reports are collected and analyzed.

The diversity of experimental procedures in those studies using protocol analysis and its implications have been discussed in Cabello and O'Hora (2002), but it is important that they are considered crucial issues because the relationship between verbal reports and task performance might well be sensitive to differences in procedure (see Austin & Delaney, 1998; Critchfield, Tucker, & Vuchinich, 1998). In spite of this diversity, however, in general these studies have demonstrated that through the use of protocol analysis, concurrent verbal behavior can be identified that is significantly related to specific task performances, which indicates its utility for the analysis of verbal behavior.

THE "SILENT DOG" METHOD

What is the silent dog?

Protocol analysis addresses some of the traditional criticisms to the use of verbal reports by measuring multiple instances of verbal behavior concurrently to the development of the experimental task, but it is not problem-free. There are a number of remaining objections that must be taken into account when using this method (Cabello & O'Hora, 2002; Hayes, 1986; Hayes, White, & Bissett, 1998): (1) It is possible that the requirement to "think aloud" introduces an extraneous variable that influences participants' behavior, so that the performances of participants who are requested to think aloud are different than those without that requirement; (2) Task performance might not require of the implication of verbal behavior, so descriptions of that performance are functionally irrelevant; (3) It is also possible that describing what participants are doing is not relevant for the ongoing task, and that these descriptions are under the control of other variables; (4) The protocol analysis is a correlational method, and thus even if verbal reports are consistently related to task performance, from a functional perspective it is not clear whether the task performance caused the verbal reports, the verbal reports caused the task performance, or both are caused by a third variable.

These objections are originated because, when using protocol analysis, behavior analysts do not ask the same questions as cognitive psychologists. According to Hayes, White and Bissett (1998), cognitive researchers are interested in determining if verbal reports reflect the underlying cognitive processes and whether these processes are the same as without these reports, while behavior analysts are interested on determining if

the task is governed by private rules, and whether the overt verbalization is functionally equivalent to the private rule. Or in other words, from a cognitive perspective the goal is determine if the content of verbal reports is formally equal to the content of cognitive processes, but from the behavioral perspective it is to demonstrate a functional equivalence without referring to the specific content of the reports.

Thus, a behavioral-analytic approach is not interested in the formal or topographical equivalence between verbal reports and private verbal behavior, but in a functional equivalence. To achieve this goal, the protocol analysis method is not enough and Hayes (1986) suggested the use of a series of methodological controls that are designed to demonstrate (a) that the performances of participants exposed to an experimental task are rule-governed, at least partially, and (b) that public verbal reports are functionally equivalent to private verbal behavior. These controls were named the silent dog method, after the Sherlock Holmes' novel "The Silver Blaze" (Doyle, 1892), in which the famous detective knew the identity of a murderer because the dog in the house had not barked.

The silent dog controls are as follows (for further information, see Barnes-Holmes, Hayes, & Dymond, 2001; Hayes, White, & Bissett, 1998):

1. No differences are found due to concurrent think-aloud: if verbal reports are functionally identical to private self-rules, then subjects are saying aloud what they are already saying privately. Thus, saying out loud what they are already saying to themselves should not affect task performance. In other words, performance on a task should be similar whether participants are required to think aloud or not.
2. Disrupting verbal behavior must affect subjects' performances: a lack of effect on performance does not necessarily mean that verbal reports are equivalent to self-rules: it could mean that they are irrelevant to the task. To avoid this problem, verbal behavior is interrupted (e.g. by asking subjects to say out loud what they thought about over the last minute). Self-rules should not be present during such disruption; therefore, if disruption affects task performance, the lack of differences cannot be attributed to the irrelevance of verbal behavior.
3. Presenting other subjects with the protocols should affect their performances in a consistent and replicable manner: the protocols should be presented to other participants about to engage in the experiment, indicating that they should consider the material. This must alter the performance of the new subjects in a consistent and replicable manner, related to the content of the protocols. This control demonstrates that the content of the protocols itself is task-relevant.

The silent dog method uses these three controls (the comparison between conditions with and without self-reports, the disruption of verbal behavior, and the replication across subjects of the effects of verbal reports) to ensure the functional similarity between public verbal reports and private self-rules. If the three controls are implemented and the expected pattern of results is obtained, then it can be concluded that: (a) the behavior is in part governed by concurrently available rules; and that (b) the lack of a difference between performance with and without concurrent talk-aloud is explained by the functional similarity of the rules present in the two conditions. That is, the self-rules formulated in the silent condition and the overt verbalizations in the talk-aloud condition are functionally the same (Barnes-Holmes, Hayes, & Dimond, 2001, p. 137).

If this pattern is not obtained, then no valid inferences can be made from the self-reports. Therefore, the proper use of the silent dog method could address the limitations about protocol analysis indicated above, establish the relevance of verbal reports through the empirical demonstration of a pattern of results, and allow for an experimental analysis of verbal behavior in which verbal reports are used as an independent variable to produce certain types of behavior (as we will discuss later).

Available empirical evidence

Despite the utility that the silent dog method could have in the analysis of verbal behavior, and despite having been described nearly 30 years ago, to date the empirical evidence is limited. Some investigations have employed some of the ideas, for example comparing the performance of participants instructed to talk aloud versus participants who completed the task in silence (Wulfert, Dougher, & Greenway 1991; Wulfert, Greenway, & Dougher, 1994), or disrupting verbal behavior during the task (Bentall, Dickins & Fox, 1993; Dickins, Bentall, & Smith, 1993; Mandell & Sheen, 1994; Sato, 2001); but only a reduced number of studies have attempted to explicitly apply the method. Table 2 summarizes these studies.

A review of the evidence shows important differences in the experimental procedures that affect how verbal reports are collected and analyzed, and also how the silent dog controls are applied, and these can be summarized as follows. First, as we also indicated about studies using protocol analysis, there is no standard procedure to collect concurrent verbal reports, with some studies following Ericsson and Simon's suggestion and some others not doing so. Second, instructions regarding the think aloud requirement and about the disruptive tasks are very different among studies, with researchers not employing the same methods. Third, although authors claim to use the silent dog method in their studies, the three methodological controls are not always applied, with some studies using only controls 1 and 2. Fourth, there is no agreement on how to apply the silent dog controls; in some cases, this is done by comparing different groups of participants, but in others, comparisons are done within-subjects. Fifth, there is a significant variability in the use of disruptive tasks to obstruct verbal behavior, with a variety of such tasks being applied (from reciting the alphabet to counting to solving math problems). Also, there is no agreement on how these tasks should be presented: during the entire experiment, or just during specific periods of time. Lastly, how to use the transcribed protocols to influence the behavior of new participants is also a problematic issue, with very different procedures being used.

Despite these differences, results consistently suggest that it is possible to apply the full set of methodological controls from the silent dog method to a number of different research areas and to participants with different verbal repertoires, and that through the demonstration of a functional equivalence between verbal reports and private rules, the role of verbal behavior and language can be better studied in human participants.

Table 2. Characteristics of published studies that have explicitly used the silent dog method. For each study, it is indicated what it was aimed at, the type of participants, how verbal reports were obtained, how the silent dog method was implemented, which controls were applied, and the main results.

	Aimed at	Training shopping abilities in children with developmental delay
	Participants	Children with delay
	Procedure for verbal reports	Specific instructions to think aloud, practice exercises. Experimenter remained in the room
Taylor & O'Reilly (1997)	Procedure for silent dog	Participants followed instructions in a role-playing situation while talking aloud, and after completed the same situation while remaining silent. Private verbal behavior was disrupted asking children to repeat a series of random numbers. Instructions used to train the shopping abilities were later used with new participants
	Applied controls	Controls 1, 2 and 3. Applied within-participants
	Main results	No significant differences were found between completing the task talking aloud or in silence. Disrupting verbal behavior affected performances, and instructions were successfully used to train new participants
	Aimed at	Examining the role of verbal behavior in a relational responding task
	Participants	Adults (students)
	Procedure for verbal reports	Not specified
Rehfeldt & Dixon (2000)	Procedure for silent dog	A first group of participants completed the task in silence, a second group was instructed to think aloud, and a third group was instructed to recite the alphabet while completing the task
	Applied controls	Controls 1 and 2. Applied between-groups
	Main results	Verbal reports did not alter task performance, as results from group 1 and 2 were equivalent. Disrupting verbal behavior in group 3 did block task performance. Verbal reports were closely related to specific task performances
	Aimed at	Studying the role of verbal behavior on the performance of participants exposed to several reinforcement schedules and on insensitivity to contingencies
	Participants	Adults (students)
Cabello, Luciano, Gómez, & Barnes-Holmes (2004)	Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter was not present in the room
	Procedure for silent dog	Four experimental conditions were used that combined completing the task in silence, completing the task while thinking aloud, and presenting several simultaneous activities to disrupt ongoing verbal behavior (from counting to solving math problems)
	Applied controls	Controls 1 and 2; control 3 was demonstrated by previous research. Applied between-groups
	Main results	Requesting participants to think aloud did not significantly alter their performances, but the simultaneous activities did so. Specific types of verbal reports were related with differential adjustment to the schedules of reinforcement

Table 2. Characteristics of published studies that have explicitly used the silent dog method. For each study, it is indicated what it was aimed at, the type of participants, how verbal reports were obtained, how the silent dog method was implemented, which controls were applied, and the main results (cont.)

	Aimed at Participants Adults (students)	Investigating the role of verbal behavior in the emergence of derived relations
	Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter was not present in the room
Cabello (2005)	Procedure for silent dog	Two experiments were conducted. In the first experiment, an equivalence task was presented using stimuli from different modalities (visual and olfactory). A first group of participants completed the task in silence, a second group were instructed to think aloud, and a third group were instructed to recite numbers and letters to disrupt verbal behavior. In the second experiment, verbal reports obtained from the think aloud group were used to influence the performances of new participants
	Applied controls	Controls 1, 2 and 3. Applied between-groups
	Main results	Criteria for the full set of silent dog controls were met. Performances during the equivalence task were rule-governed and verbal reports in which participants established multiple relations among stimuli, were related to better performances. These rules also produced better results when used as instructions for new participants exposed to the task
	Aimed at Participants Adults (workers)	Studying the relation among rules and behavioral safety at the workplace
	Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter was not present in the room
Alvero & Austin (2006)	Procedure for silent dog	Two experiments were conducted. In the first experiment, a task that simulated a job situation was used, and the performances of participants with and without verbal reports were compared, as well as during the presentation of a disruptive task. In the second experiment, the impact of verbal reports on the performance of new participants was analyzed
	Applied controls	Controls 1, 2 and 3. Applied between-groups
	Main results	Criteria for the three silent dog controls were met. Results demonstrated a functional relation between verbal utterances related to behavioral safety at the workplace, and an actual increment in effective safety behaviors
	Aimed at Participants Two children with autism	Studying the role of verbal regulation on the non verbal behavior of children with autism
	Procedure for verbal reports	Specific instructions to think aloud, practice exercises, reminders to think out aloud. Experimenter remained in the room
Arntzen, Halstadro, & Halstadro (2009)	Procedure for silent dog	The first participant completed several computerized tasks both silently and thinking aloud, and also while presented with disruptive tasks (solving math problems). Verbal reports were collected and later used to train a second participant
	Applied controls	Controls 1, 2 and 3. Applied both within-participants and between-participants
	Main results	Criteria for the three silent dog controls were met. The performance of both participants appeared to be under the control of self-rules, and verbal reports from the first participant were effective to train the computerized tasks for the second participant

WHAT IS OF IMPORTANCE IN PROTOCOL ANALYSIS AND THE SILENT DOG METHOD

Several times in this paper we have argued that protocol analysis, coupled with the methodological controls from the silent dog method, addresses most criticisms about the validity of verbal reports in the experimental analysis of behavior. Using these tools, the researcher can ensure, empirically, the functional correspondence between private verbal rules that participants could be following and what they publicly report, because the validation of verbal reports as data is done through the demonstration of a specific pattern of results (i.e., no differences due to the “think aloud” requirement, alterations in task performance when verbal behavior is disrupted, and the replication of the effect of verbal reports in new participants). This type of validation has been termed “empirical validation” and is to be preferred from a scientific point of view, because it does not rely on theoretical elaborations or on implicit assumptions (Critchfield & Epting, 1998; Critchfield, Tucker, & Vuchinich, 1998).

This is an idea worth remembering when examining the current literature of studies that have used protocol analysis and the silent dog controls, because most times their findings are in accordance with previous research and it could be argued that nothing new has been discovered. What is new, and relevant, are not the specific results or the specific relations between verbal report and nonverbal performances that were found, but that researchers used sound methods to demonstrate that the performances of participants were rule-governed by private rules, that verbal reports were functionally equivalent to these private rules, that the content of these reports was task-relevant and that inferences drawn from the participants’ verbal behavior were adequate and empirically supported. This is the power of the two methods described above.

Another relevant issue that must be kept in mind is that through these two methods, the question that the behavior analyst must formulate is “Are verbal reports functionally equivalent to private rules?”, and not “Is the content of the verbal reports the same as of the private rules?”. As discussed previously, the goal of the silent dog method is not to make public the same behavior that is ongoing privately, because it is not necessary that verbal reports exactly reflect the private rules; it is sufficient that both verbal reports and private rules are demonstrated to be under control of the same variables, and thus that they are functionally equivalent (Austin & Delaney, 1998; Cabello & O’Hora, 2002; Barnes-Holmes, Hayes, & Dymond, 2001; Hayes, White, & Bissett, 1998). This is an important difference from how cognitive psychologists employ protocol analysis, in which for the verbal reports to be relevant, they must have the same content of cognitive processes. This is impossible to demonstrate because private mental processes can not be externally and empirically verified; thus, cognitive studies have to rely on “conceptual validation”. Indeed, astute readers could argue that behavior analysis is not actually using the protocol analysis technique as was proposed by Ericsson and Simon (1980), but something similar because although the recommendations regarding “thinking aloud” and data analysis are used, the goals are fundamentally different and additional methodological controls should be added to ensure data validity.

Furthermore, it must be remembered that the silent dog method should be used (at least partially) when using the protocol analysis technique, because in what probably

is the most important strength of the recommended methods, the procedure permits an experimental analysis of verbal rules, and not merely a correlational one. A review of most studies based on verbal reports (including those in which only protocol analysis is applied) shows that they are of correlational nature, and that even if verbal reports are significantly related to specific task performances, from a functional point of view it is not clear whether (a) task performances caused the verbal reports, (b) the verbal reports caused task performances, or (c) both are caused by a third variable. Therefore, a satisfactory explanation must specify and demonstrate the current and historical contexts that are responsible for this relation between behaviors (Barnes, 1989; Critchfield, Tucker, & Vuchinich, 1998; Hayes, 1986; Hayes & Brownstein, 1986; Hayes, White, & Bissett, 1998; Hinline & Wanchisen, 1989; Luciano, 1992, 1993; Shimoff, 1986). In the case of studies using the silent dog method, although a correlational strategy is used at a first stage, verbal reports are used to affect the performance of new subjects. Self-reports are now the independent variable and are used to produce specific patterns of responding; therefore, verbal behavior is brought under explicit and replicable experimental control (Cabello & O'Hora, 2002).

CONCLUSIONS

The main goal of this paper was to analyze the role of verbal reports in the experimental analysis of behavior, and to determine if what participants say in the laboratory can be considered valid data for the study of human behavior. As we have seen, the interest on these types of measures in behavior analysis was reduced for many years, and the few studies that collected verbal reports did not conform a consistent line of work. During the 80s, however, the work on areas such as stimulus equivalence (Sidman & Tailby, 1982; Sidman, 1994), insensitivity to contingencies (Gómez, 1996; Madden, Chase, & Joyce, 1998) and rule-governed behavior (Hayes, 1989) produced a shift to the study of behavior in human adults that which in turn increased the interest in experimental procedures that collected verbal reports; however, the validity of these reports was critiqued and taken with caution.

Despite these critiques, and contrary to what some could argue, behavior analysis never rejected verbal reports. Shortcomings and limitations that could affect their use were always kept at a methodological level and as part of the traditional emphasis on sound experimental procedures. Therefore, it was stated that for verbal reports to be useful in the analysis of human verbal behavior, their validity should be demonstrated and any inference drawn from them should be supported by empirical evidences (Critchfield, Tucker, & Vuchinich, 1998; Perone, 1988; Shimoff, 1986).

We have proposed here that over the last three decades, two methods have been proposed that ensure that verbal behavior is relevant for an experimental task and that verbal reports obtained concurrently are functionally similar to private verbal behavior. These tools are the protocol analysis and the silent dog method, of which the latter is probably the most relevant from a behavioral perspective (Hayes, 1986; Hayes, White, & Bissett, 1998), because through a series of methodological controls it addresses most criticisms raised against verbal reports, and it allows for an experimental analysis of

verbal behavior based on empirical, valid and contrastable data. In light of this, we believe that behavior analysis can use verbal reports at the same level of confidence than other well established and widely used measures, given that their collection and analysis are performed carefully, as is in the case of protocol analysis and the silent dog method.

It is legitimate, then, to question the role of other procedures based on reports from participants, such as post-experimental reports, written questionnaires or pen-and-paper scales. In these cases the functional equivalence between those verbal reports and their referents cannot be adequately established, any conclusion obtained from them should be taken with precaution and always remain at a correlational level. We do not affirm that verbal reports obtained through these other procedures should be dismissed, but that it must be recognized that, at the moment, our experimental technology does not allow for their empirical validation, and therefore we do not know whether they reflect relevant, private verbal behavior. As an example, recent works conducted using the Implicit Relational Assessment Procedure (e.g. Cullen, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009; Maroto, Hernández, & Rodríguez, 2015; Power, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009; Vahey, Boles, & Barnes-Holmes, 2010) have demonstrated that post-experimental reports in socially sensitive areas do not correspond to the verbal and relational repertoire of participants, and can lead to inadequate inferences. Of course, this does not mean that protocol analysis and the silent dog method are a solution for all problems or even that they can be used in all research areas, but that they are our best, more powerful tools of analysis today.

Lastly, we would like to acknowledge the necessity for additional research on both tools, to further develop them; more specifically, there are three areas in which new research is urgently needed. First, it is necessary to conduct studies that clarify the influence of the different procedures that have been used in the literature, so that we better understand the best way to instruct participants, collect verbal reports, or implement the silent dog controls. For example Plancarte, Moreno, Hickman, Arroyo, & Cepeda (2013) conducted a recent study in which they demonstrated that differential effects were found depending on whether self-reports were obtained concurrently or retrospectively. Further studies should directly compare these different procedures and measure how they impact the verbal reports produced and the relation of these reports with ongoing private verbal behavior.

Second, it is necessary to start a line of work that applies both protocol analysis and the silent dog method systematically and to more diverse research topics. To date, most of the available studies are isolated efforts and with no continuity. Indeed, a pattern that is repeated among most authors is to conduct an initial study that is regarded as “very interesting”, but given the difficulty of the analysis, the effort needed to collect, transcribe, codify and analyze verbal reports, and the effort to apply adequate experimental controls, they soon turn to less demanding studies.

Third, although some studies have done so, it is necessary to further apply these procedures to children. Because their verbal repertoire is less sophisticated than that of adults who typically participate in experiments, the conclusions about the role of private verbal behavior on human functioning can be more interesting (Luciano,

Barnes-Holmes, & Barnes-Holmes, 2001). This is especially relevant given that, from a cognitive perspective, a large number of studies using verbal reports are conducted with children (e.g., Berk, 1999; Winsler, Fernyhough, & Montero, 2009).

To conclude, at the beginning of this paper we have stated that human participants, during experimental tasks, show a variety of verbal behavior such as counting, describing or reasoning. Although the analysis of verbal reports has been an object of debate within behavior analysis, currently we have research tools that allow for such analysis. It is necessary to use these tools in new experimental studies, because not paying attention to what participants say (or can say) about their own behavior, and to its relation with other behaviors, is a mistake we cannot afford to make as a scientific discipline.

REFERENCES

- Alvero A & Austin J (2006). An implementation of protocol analysis and the silent dog method in the area of behavioral safety. *Analysis of Verbal Behavior*, 22, 61-79.
- Arntzen E, Halstadro L & Halstadro M (2009). The silent dog method: analyzing the impact of self-generated rules when teaching different computer chains to boys with autism. *The Analysis of Verbal Behavior*, 25, 51-66.
- Austin J & Delaney PF (1998). Protocol analysis as a tool for behavior analysis. *The Analysis of Verbal Behavior*, 15, 41-56.
- Barnes D (1989). Behavior-behavior analysis, human schedule performance and radical behaviorism. *The Psychological Record*, 39, 339-350.
- Barnes-Holmes D, Hayes SC, & Dymond S (2001). Self and self-directed rules. In SC Hayes, D Barnes-Holmes & B Roche (Eds). *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pages 119-139). New York: Plenum Press.
- Bentall RP Dickins DW & Fox SRA (1993). Naming and equivalence: response latencies for emergent relations. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology*, 46B, 187-214.
- Bentall RP, Lowe CF, & Beasty A (1985). The role of verbal behavior in human learning: I. Developmental differences. *Journal of the Experimental Analysis of Behavior*, 43, 165-181.
- Berk L (1999). Children's private speech: An overview of theory and the status of research. In P Lloyd & C Fernyhough (Eds). *Vygotsky: Thought and Language*. Florence: Routledge.
- Bulbrook ME (1932). An experimental inquiry into the existence and nature of "insight". *American Journal of Psychology*, 44, 409-453.
- Buskist WF & Miller HL (1982). The analysis of human operant behavior: A brief census of the literature, 1958-1981. *The Behavior Analyst*, 5, 137-141.
- Cabello F (2005). *Aprendizaje relacional de distintas modalidades estímulares: implicación de la conducta verbal [Relational learning of different stimulus modalities: Implication of verbal behavior]*. Unpublished doctoral dissertation, Universidad de Almería.
- Cabello F, Luciano MC, Gómez I, & Barnes-Holmes D (2004). Human Schedule performance, self-rules and the silent dog strategy. *The Psychological Record*, 54, 405-422.
- Cabello F & O'Hora D (2002). Addressing the limitations of protocol analysis in the study of complex human behavior. *International Journal of Psychology and Psychological Therapy*, 2, 115-130.
- Catania AC & Shimoff E (1998). The experimental analysis of verbal behavior. *The Analysis of Verbal Behavior*, 15, 97-100.
- Conrad F, Blair J, & Tracy E (1999). *Verbal reports are data! A theoretical approach to cognitive interviews*. Proceedings of the Federal Committee on Statistical Methodology Research Conference, Arlington (pp. 11-20).
- Critchfield TS & Epting L K (1998). The trouble with babies and the value of bathwater: complexities in the use of verbal reports as data. *The Analysis of Verbal Behavior*, 15, 65-74.
- Critchfield TS, Tucker JA, & Vuchinich RE (1998). Self-report methods. In K Lattal & M Perone (Eds).

- Handbook of research methods in human operant behavior* (pages 435-470). New York: Plenum Press.
- Crutcher RJ (1994). Telling what we know: the use of verbal report methodologies in psychological research. *Psychological Science*, 5, 241-245.
- Cullen C, Barnes-Holmes D, Barnes-Holmes Y, & Stewart I (2009). The Implicit Relational Assessment Procedure (IRAP) and the malleability of ageist attitudes. *The Psychological Record*, 59, 591-620.
- Dickins DW (2015). Vocalizing phonologically correct non-word stimuli during equivalence class formation. *European Journal of Behavior Analysis*, 16, 248-278.
- Dickins DW, Bentall RP, & Smith AB (1993). The role of individual stimulus names in the emergence of equivalence relations: the effects of interpolated paired-associates training of discordant associations between names. *The Psychological Record*, 43, 713-724.
- Dixon MR & Hayes LJ (1998). Effects of differing instructional histories on the resurgence of rule-following. *The Psychological Record*, 48, 275-292.
- Doyle AC (1892). *The memoirs of Sherlock Holmes*. Reprinted as *Sherlock Holmes: Complete novels and stories* (vol. 1). New-York: Bantam Books.
- Dube WV, Green G, & Serna RW (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, 59, 103-114.
- Dugdale NA (1988). *The role of naming in stimulus equivalence: Differences between humans and animals*. Unpublished doctoral dissertation, University of Wales.
- Duncker K (1945). On problem-solving. *Psychological Monographs*, 58, 5, whole number.
- Dymond S, Roche B, & Barnes-Holmes D (2003). The continuity strategy, human behavior, and behavior analysis. *The Psychological Record*, 53, 333-347.
- Ericsson KA (2003). Valid and non-reactive verbalization of thoughts during performance of tasks: towards a solution to the central problems of introspection as a source of scientific data. *Journal of Consciousness Studies*, 10, 1-18.
- Ericsson KA & Simon HA (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Ericsson KA & Simon HA (1993). *Protocol analysis: Verbal reports as data* (Rev. Ed.). Cambridge, MA: MIT Press.
- Fukui I (2002). A review of recent topic in the experimental analysis of human behavior. *Japanese Journal of Experimental Analysis of Behavior*, 34, 12-32.
- García Y & Rehfeldt RA (2008). The effects of common names and common FR responses on the emergence of stimulus equivalence classes. *European Journal of Behavior Analysis*, 9, 99-120.
- Gómez I (1996). *Investigación sobre el fenómeno de insensibilidad a las contingencias [A research of the phenomenon of insensitivity to contingencies]*. Unpublished doctoral dissertation, Universidad de Almería.
- Hayes SC (1986). The case of the silent dog-verbal reports and the analysis of rules: a review of Ericsson and Simon's Protocol Analysis: Verbal reports as data. *Journal of the Experimental Analysis of Behavior*, 45, 351-363.
- Hayes SC (1987). Upward and downward continuity: It's time to change our strategic assumptions. *Behavior Analysis*, 22, 3-6.
- Hayes SC (1989). *Rule-governed behavior: Cognition, contingencies and instructional control*. New York, Plenum Press.
- Hayes SC & Brownstein AJ (1986). Mentalism, behavior-behavior relations and a behavior analytic view of the purposes of science. *The Behavior Analyst*, 9, 175-190.
- Hayes SC, White D, & Bissett RT (1998). Protocol analysis and the silent dog method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior*, 15, 57-63.
- Hineline PN & Wanchisen BA (1989). Correlated hypothesizing and the distinction between contingency-shaped and rule-governed behavior. In SC Hayes (Ed.) *Rule-governed behavior: Cognition, contingencies and instructional control* (pages 221-268). New York, Plenum Press.
- Horne PJ & Lowe CF (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65, 185-241.
- Hyten C & Reilly M (1992). The renaissance of the experimental analysis of human behavior. *The Behavior Analyst*, 15, 109-114.

- Lowe CF, Harzem P, & Hughes S (1978). Determinants of operant behavior in humans: Some differences from animals. *Quarterly Journal of Experimental Psychology*, *30*, 373-386.
- Luciano C (1992). Significados aplicados de los tópicos de investigación básica conocidos como relaciones de equivalencia, decir-hacer y sensibilidad e insensibilidad a las contingencias [Applied meaning of basic research topics known as equivalence relations, say-do and sensitivity and insensitivity to contingencies]. *Análisis y Modificación de Conducta*, *18*, 805-859.
- Luciano C (1993). La conducta verbal a la luz de recientes investigaciones. Su papel sobre otras conductas verbales y no verbales [Verbal behavior in light of recent research. Its role on other verbal and nonverbal behaviors]. *Psicothema*, *5*, 351-374.
- Luciano C, Barnes-Holmes Y, & Barnes-Holmes D (2001). Early verbal developmental history and equivalence relations. *International Journal of Psychology and Psychological Therapy*, *1*, 137-149.
- Madden GJ, Chase PN, & Joyce JH (1998). Making sense of sensitivity in the human operant literature. *The Behavior Analyst*, *21*, 1-12.
- Mandell C & Sheen V (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Processes*, *32*, 29-46.
- Maroto P, Hernández M & Rodríguez M (2015). Assessment of implicit anti-fat and pro-slim attitudes in young women using the Implicit Relational Assessment Procedure (IRAP). *International Journal of Psychology and Psychological Therapy*, *15*, 17-32.
- Miller NE (1985). The value of behavioral research on animals. *American Psychologist*, *40*, 423-440.
- Moreno D, Ribes E, & Martínez C (1994). Evaluación experimental de la interacción entre el tipo de pruebas de transferencia y la retroalimentación en una tarea de discriminación condicional bajo aprendizaje observacional [Experimental assessment of the interaction between transfer tests type and feedback in a conditional discrimination task under observational learning]. *Revista Latina de Pensamiento y Lenguaje*, *2*, 245-266.
- Navarick DJ, Bernstein DJ, & Fantino E (1990). The experimental analysis of human behavior. *Journal of the Experimental Analysis of Behavior*, *54*, 123-128.
- Newell A & Simon HA (1972). *Human problem solving*. Englewood Cliffs, Prentice-Hall.
- Normand MP (2001). *Collecting concurrent verbal reports: Some considerations and methodological refinements*. Paper presented at the Association for Behavior Analysis Meeting, New Orleans.
- Normand MP & Bailey JS (2006). The effects of celeration lines on visual analysis. *Behavior Modification*, *30*, 295-314.
- Perone, M. (1988). Laboratory lore and research practices in the experimental analysis of human behavior: use and abuse of subjects' verbal reports. *The Behavior Analyst*, *11*, 71-15.
- Plancarte P, Moreno D, Hickman H, Arroyo R, & Cepeda ML (2013). Use of analysis of verbal protocols in the study of complex human behavior. In C García, V Corral Verdugo & D Moreno (Eds.). *Recent Hispanic research on sustainable behavior and interbehavioral psychology* (pages 159-178). Hauppauge, NY: Nova Science Publishers.
- Potter B (1999). Some additional considerations of protocol analyses. *The Analysis of Verbal Behavior*, *16*, 57-61.
- Potter B, Huber S, & Michael J (1997). The role of mediating verbal behavior in selection-based responding. *The Analysis of Verbal Behavior*, *14*, 41-56.
- Power PM, Barnes-Holmes D, Barnes-Holmes Y, & Stewart I (2009). The Implicit Relational Assessment Procedure (IRAP) as a measure of implicit relative preferences: A first study. *The Psychological Record*, *59*, 621-640.
- Rehfeldt RA & Dixon MR (2000). Investigating the relation between self-talk and emergent stimulus relations. *Experimental Analysis of Human Behavior Bulletin*, *18*, 28-29.
- Rehfeldt RA, Dixon MR, Hayes LJ, & Steele A (1998). Stimulus equivalence and the blocking effect. *The Psychological Record*, *48*, 647-664.
- Rehfeldt RA & Hayes LJ (2000). The long-term retention of generalized equivalence classes. *The Psychological Record*, *50*, 405-428.
- Sato T (2001). The function of common naming and establishment of stimulus equivalence. *Japanese Journal of Behavior Analysis*, *16*, 2-21.
- Schlinger H (1998). Editorial. *Analysis of Verbal Behavior*, *15*, 1-2.

- Shimoff E (1986). Post-session verbal reports and the experimental analysis of behavior. *The Analysis of Verbal Behavior*, 4, 19-22.
- Sidman M (1994). *Stimulus equivalence: A research story*. Boston: Author's Cooperative.
- Sidman M & Tailby W (1982). Conditional discrimination versus matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5-22.
- Skinner BF (1957). *Verbal behavior*. New York: Appleton-Century-Crofts
- Skinner BF (1984). An operant analysis of problem solving. *Behavioral and brain sciences*, 7, 583-613
- Vahey N, Boles S, & Barnes-Holmes D (2010). Measuring adolescents' smoking-related social identity preferences with the Implicit Relational Assessment Procedure (IRAP) for the first time: A starting point that explains later IRAP evolutions. *International Journal of Psychology and Psychological Therapy*, 10, 453-474.
- Watson JB (1920). Is thinking merely the action of language mechanisms? *British Journal of Psychology*, 11, 87-104.
- Winsler, A, Fernyhough C, & Montero I (2009). *Private speech, executive functioning, and the development of verbal self-regulation*. New York: Cambridge University Press.
- Wulfert E, Dougher MJ, & Greenway DE (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *Journal of the Experimental Analysis of Behavior*, 56, 489-504.
- Wulfert E, Greenway DE, & Dougher MJ (1994). Third-order equivalence classes. *The Psychological Record*, 44, 411-439.

Received, February 13, 2016
Final Acceptance, April 1, 2016