Contrasting Alternatives Rating Scale

Ibrahim A. Halloun

Authentic assessment requires a variety of tools to paint a comprehensive picture of students' profiles, and to subsequently allow teachers make informed and constructive decisions about learning and instruction. These tools are about both the core-dimension and the meta-dimension of a given course. The *core*-dimension is discipline specific, and often course specific. It includes *content* and *process* knowledge that pertains to a given discipline or course. Content knowledge consists of all *conceptions* (concepts, laws and other theoretical statements) and especially *conceptual models* of a particular scientific theory that students are expected to develop in a particular course. Process knowledge consists of all tools, norms, rules and guidelines, especially of conceptual nature, needed for a student to meaningfully construct and deploy the target content knowledge. The *meta*-dimension is generic and it may cover a variety of courses in a given discipline. It includes general *dispositions* (e.g., beliefs, attitudes, affects) that pertain to the nature of the discipline itself, as viewed in scientific and educational contexts (i.e., from philosophical and pedagogical perspectives), as well as to generic aspects of science and science education. It also includes student learning styles and faculty teaching styles. The Contrasting Alternatives rating scale (CArs) is most appropriate for assessing the meta-dimension, but it may also be used in the assessment of the core-dimension.

I developed CArs originally for the Views About Science Survey (VASS) in the mid nineties. VASS is an instrument to assess student dispositions about knowing and learning science. More specifically, VASS addresses scientific inquiry and the structure and validity of scientific theory, as well as learning styles with a focus on reflective thinking, and the relevance and intellectual feasibility of science courses. VASS started as an open-ended questionnaire with the intention to turn it eventually into a closed format so that it becomes feasible to administer to large student populations. Multiple-choice and Likert formats were then considered, and tried, in a later stage. Soon, however, it became evident that such formats were not suitable for VASS. In fact, research then indicated, and it still does, that these formats, and especially the Likert scale, exhibit insuperable validity and reliability problems when used in surveying students' dispositions and learning styles. I then came up with the Contrasting Alternatives rating scale (CArs), tested this survey format and validated it for VASS (*references in my VASS publications)*.

My physics course covers:

(a) abstract themes.

1	**2**	**3**	**4**	**5**
a >> b	a > b	a = b	b > a	b >> a

(b) practical applications.

In the case of the example above, the five choices would mean the following:

1. My physics course covers *mostly* abstract themes and *rarely* any practical applications.
2. My physics course covers *more* abstract themes *than* practical applications.
3. My physics course covers *as much* abstract themes *as* practical applications.
4. My physics course covers *more* practical applications *than* abstract themes.
5. My physics course covers *mostly* practical applications and *rarely* any abstract themes.

Figure 1. A CArs item given for illustration in the physics form of VASS.

Over the years, CArs evolved from an eight-point rating scale to a five-point rating scale. In its current form, a CArs item presents two viewpoints about a particular issue that respondents need to contrast on a five-point scale. As outlined in Figure 1, a respondent may express a position that leans toward one viewpoint more than the other, or that may favor both equally. In the following is an outline of some major advantages that this rating scale offers over other formats, especially open-ended and Likert scale.

1. *Open-ended questions and CArs*

Open-ended questions are often open to a wide variety of interpretations by respondents as well as by researchers. They can be misleading, especially when respondents' priorities or value judgments are not the same as those of the concerned researcher, which is often the case when assessing the meta-dimension of a given course. When VASS was first administered in open-ended format, students were asked in one of the questions to tell what is the first thing they do in solving a physics problem. The student in Figure 2 replied by writing that he starts by looking for the appropriate *formula*. When interviewed, it became evident that the first thing this student actually does in solving a physics problem is *drawing diagrams*. However, this procedure seemed so trivial to him that he thought it was not worth mentioning in his written response.

The risk of interpretation mismatch between surveyor and respondent that we run with open-ended questions (as well as with other traditional formats) is resolved in CArs where the surveyor's position can be easily and explicitly contrasted with a distinct popular position. This latter position should of course come from open-ended questionnaires coupled with interviews to clear up any possible mismatch like the one in Figure 2. This is actually what we have done with VASS, and this is how we ended, for example, asking the question of Figure 2 in a way to contrast between the search for formulas and the drawing of diagrams as one begins solving a physics problem (§ 2).

I: Describe what you normally do when solving a physics problem. List all steps you often follow, in order.

S: First step in any problem would be to read the problem and list what's given and what you need, variables or what not. And the next step would be to find formulas that include these variables. And then, the third would be to solve for the unknowns. That's basically it.

I: So this would be an algorithm you would work through in any kind of problem?

S: Basically, I would agree. It's a basic general, general outline of how to solve a problem.

I: Do you ever consider drawing some kind of a diagram?

S: Uh-huh... I'd consider that helpful, yeah, I'd probably include that in step one. Draw, label, find out what you have and don't have.

I: So that becomes then, your first step.

S: Uh-huh.

I: Would that be true for any kind of problem?

S: Visualization helps a lot. I would say it would be a good step to try in any problem. If you can't visualize it, I wouldn't try to draw it. Yeah, I would agree that would have to be helpful for any kind of problem.

I: Do you usually do it?

S: Do I do it? Usually yes. It's almost asked of us in physics class: force diagrams, free body diagrams. I would say they're probably most helpful. I would say, yeah.

Figure 2. Excerpts from a student interview about the first thing he does when solving a physics problem.

2. *Likert rating scale and CArs*

The interpretation mismatch just discussed is common to a wide range of formats including, especially, the Likert rating scale. Two respondents may express opposite positions on a Likert item for the same reason, or the same position for contradictory reasons. When presented in a CArs format, respondents are focused on the context within which they need to answer a given question, and so are researchers in interpreting responses.

In 2000, I conducted a limited study involving 230 physics students to compare CArs to Likert scale in the context of some VASS items. I then devised a paper-and-pencil instrument consisting of two sets of questions. The first set of questions consisted of VASS items asked in a Likert format. The second set consisted of the same items asked in CArs format. The two alternatives associated with each CArs question were provided in two separate questions in the Likert questionnaire. For instance, the question addressed in Figure 2 was first presented in the following two Likert statements:

> *The first thing I do when solving a physics problem is to search for formulas that relate givens to unknowns.*

Then later in the test:

> *The first thing I do when solving a physics problem is to represent the situation with sketches and drawings.*

In this and all other Likert items, respondents were asked to express their position on a 5-point rating scale ranging from "Strongly Agree" to "Strongly Disagree", with a neutral position in the middle. Once they have completed and turned in the Likert survey, participants were given a CArs survey dealing with the same items. For example, the two Likert statements above were presented in a single CArs question framed as follows:

> *The first thing I do when solving a physics problem is:*
> *(a) represent the situation with sketches and drawings.*
> *(b) search for formulas that relate givens to unknowns.*

and respondents were asked to contrast the two alternatives on the 5-point rating scale shown in Figure 1.

Respondents systematically expressed contradictory positions on any two Likert statements dealing with the same issue. For example, in the case of the two statements above, 74% of respondents sided with "sketches and drawings" (18% did not) and 76% sided with "formulas" (14% did not). When asked the same question in CArs format, 31% were inclined "mostly" or "more" toward alternative (a), and 39% favored equally both alternatives. The same pattern of contradictions within Likert items, and discrepancies between these items and CArs items, were observed throughout all items of the study (Halloun, 2001).

The study showed that the Likert rating scale can be misleading, even when all possible viewpoints pertaining to a given issue are offered in separate questions, and that the CArs format is significantly more reliable. In the CArs format, a person's response is internally normalized; one's position has to be stated for one alternative by comparison to the contrasting one and not in the absolute sense. This normalized comparison is not made possible with the Likert format (or other formats) where there is no point of reference relative to which one can state his/her position.

3. *Multiple choice and CArs*

Some may argue that CArs could be substituted with a multiple-choice format whereby a respondent may choose either alternative, both or none. This format is limited by comparison to CArs. When a respondent picks both alternatives in a multiple-choice format, there is no way to figure out whether s/he favors the two equally or to different degrees. In fact, our experience has suggested that this is a critical option that students and faculty would like to have in instruments dealing with the meta-dimension of any course. As for the "none" option, the two alternatives in a CArs item are chosen from actual viewpoints held by experts and novices. They are practically so popular that hardly any respondent would feel that none of the two alternatives matches his/her viewpoint. In fact when we started with CArs as an eight-point rating scale, "none" was an option. VASS and CArs have evolved since to a point that hardly any respondent would pick that option. That is why it was left out.

Aside from the limited study comparing CArs to Likert, CArs was established through continuous analysis of VASS that was administered to tens of thousand students in various countries around the world. By all classical statistical standards, VASS was shown to be a valid and reliable instrument. For example, Cronbach Alpha has often come up to over .80, which, by classical standards of reliability, is an indication of a significantly high internal consistency. Faculty and researchers who have used this instrument for assessing student dispositions and evaluating instruction have consistently expressed their satisfaction with the CArs format.