

Investigation of AP Statistics Students' Understanding of Technical Terminology with Possible Lexical Ambiguities

Douglas Whitaker
Steven Foti
Tim Jacobbe

Introduction

- A review of the literature revealed dozens of terms used in statistics with non-statistical meanings
 - Kaplan, Fisher, and Rogness (2009, 2010)
 - Kaplan, Rogness, and Fisher (2012, 2014)
 - Lavy and Mashiach-Eizenberg (2009)
 - Pierce and Fontaine (2009)
 - Watson and Kelly (2008)
- These lexical ambiguities can hinder instruction
- Identified ambiguities can suggest instructional changes (e.g., substituting a different term)
- Current literature focuses on college students
- Experience and discussions suggested more words
- Examples of words (Kaplan et al. 2009, 2010):
 - Association, average, bias, blocking, center, confidence, control, correlation, distribution, error...*
 - Some words studied in-depth, others just identified

Methods

- Before instruction, data on over a dozen words were collected in an AP Statistics course (maximum N is 22).
- Six words chosen for further data collection:
 - Arbitrary, model, random, range, uniform, variable*
 - Of these, *random* is the only word of these that has been extensively studied in the literature.
- Data were transcribed and coded.
 - Initial codes were created for each distinct concept.
 - After the initial coding, related codes were collapsed into more general final codes.
 - All data were then re-coded using the final codes.

Instrument

- Before** instruction data collection example:

Instructions: For each underlined word, write a **definition** in your own words and use the word in a **sentence**.

Variable

Definition:

Sentence:

- After** instruction data collection example:

Define and use in a sentence the word variable as you use it in **daily life**.

Definition:

Sentence:

Define and use in a sentence the word variable as it is used in **statistics**.

Definition:

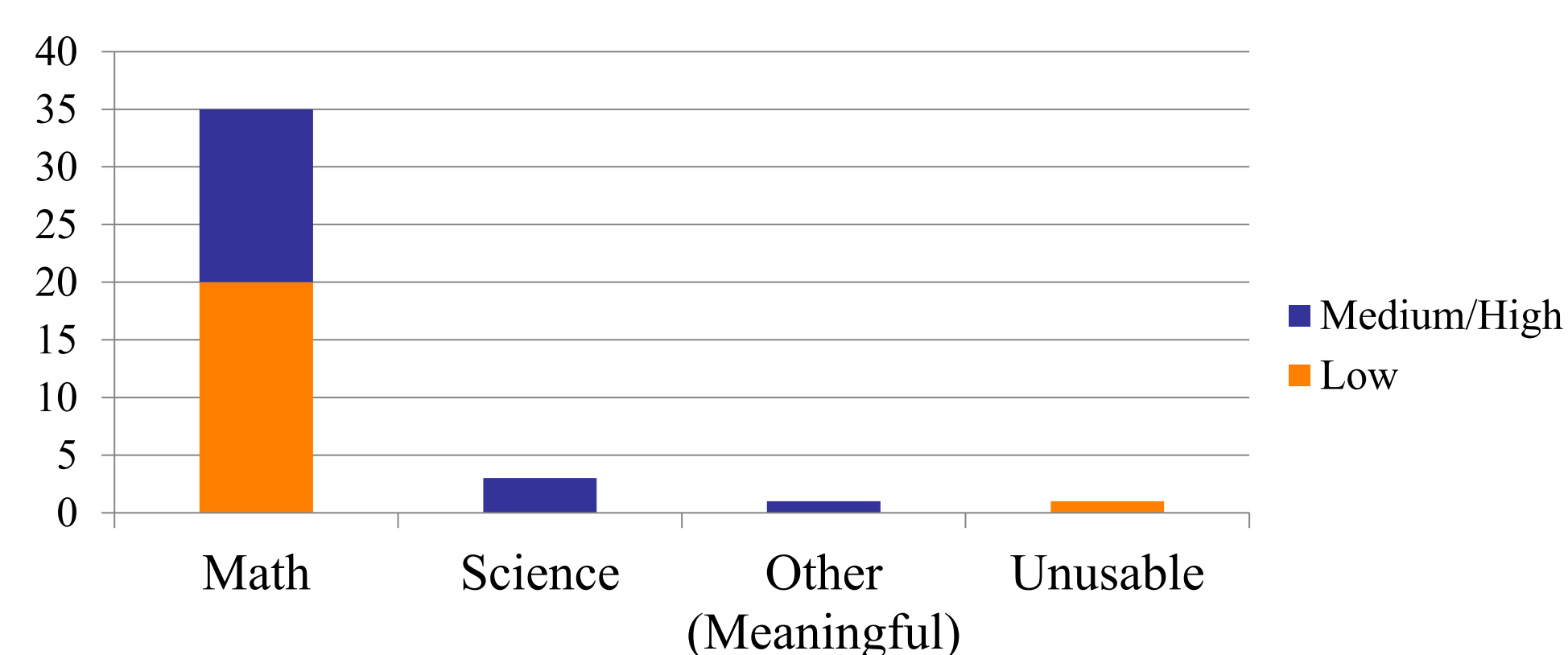
Sentence:

Example Codes for Variable

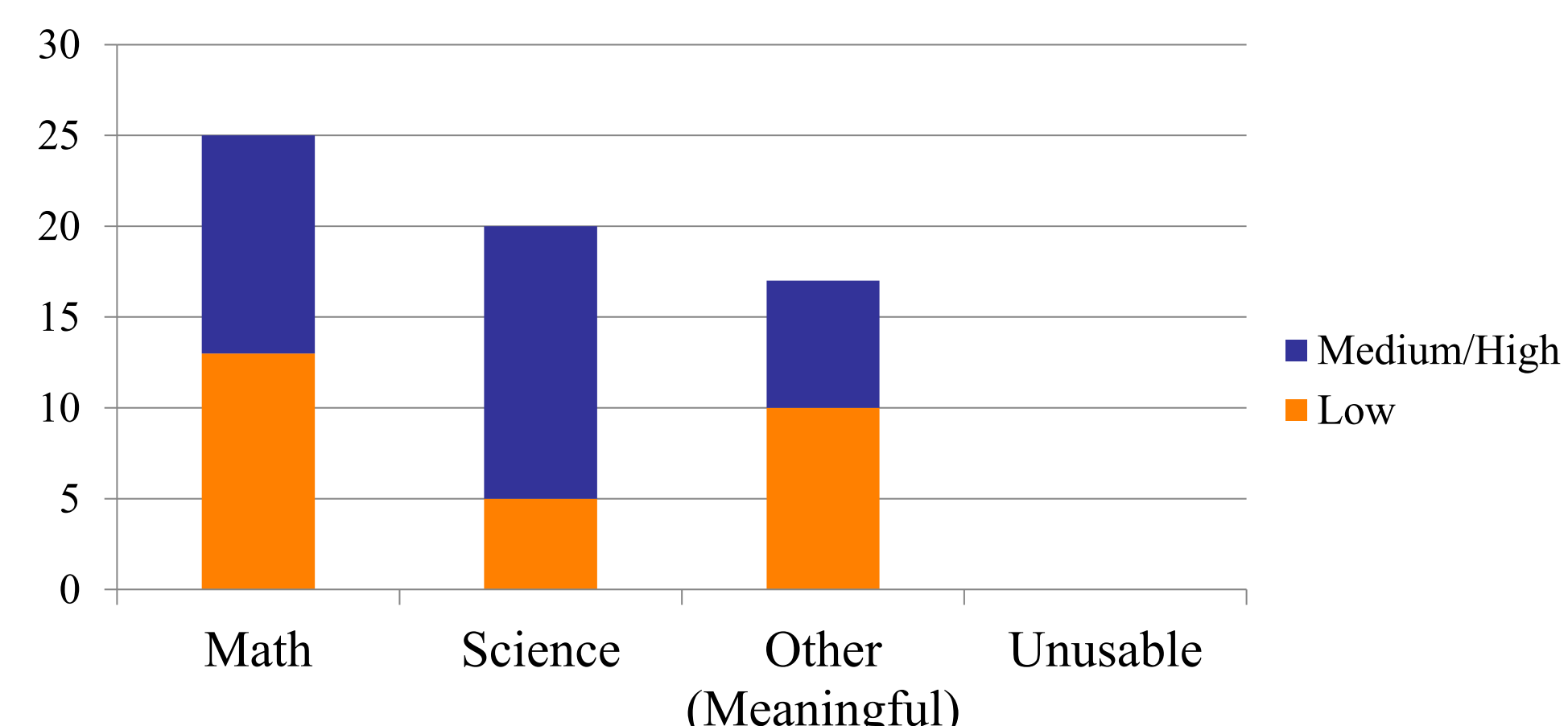
- a letter used to take the place of a number in an equation* [Math, High]
- The variable x was trying to be solved.* [Math, Medium]
- Dependent and independent variables are used to show how one variable affects another.* [Science, High]
- something that affects an outcome* [Science, Medium]
- There are a lot of variables in this situation* [Science, Low]
- The results had a large range and they were highly variable.* [Other (Meaningful), Medium]
- The data was variable.* [Other (Meaningful), Medium]
- Dang look at how variable that sandwich is.* [Unusable, Low]

Distribution of Codes for Variable

Before Instruction



After Instruction



Coding

Responses coded for two features: content and quality

Content codes for Variable are:

- Math, Science, Other (Meaningful), Unusable

Content codes for Range are:

- Formula, Variability, Other (Meaningful), Unusable

Quality codes (for both) are:

- Low, Medium, High

Preliminary Findings

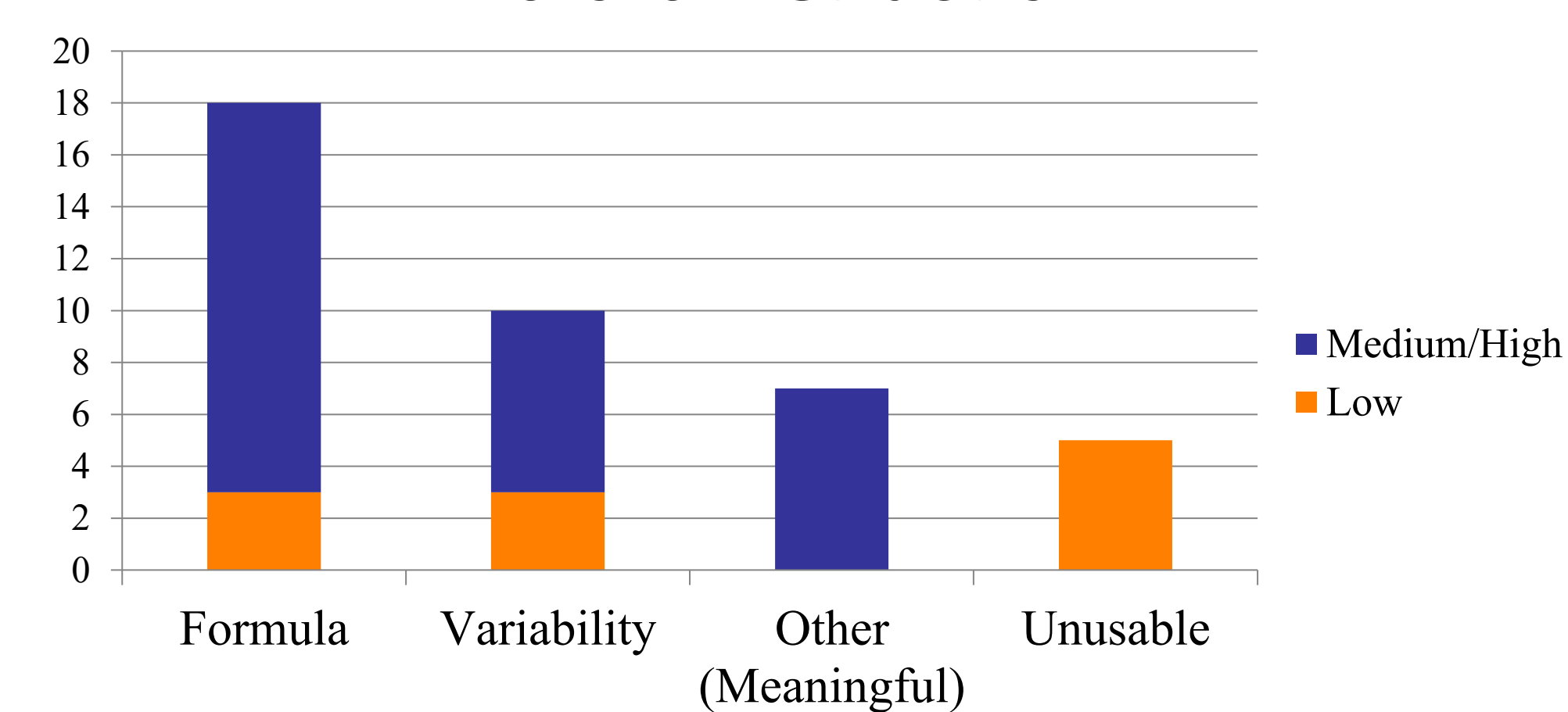
- Lexical ambiguity for many statistical terms is apparent.
- Ambiguities come from many sources, notably
 - colloquial meanings and
 - other technical terms.
- Statistics is viewed by many students as a mathematics course. These students look to meanings learned in other mathematics courses when encountering statistical terms.
- Traditional statistics coursework can improve student conceptions.
- Students' conceptions of words – even non-technical terms – may be non-standard.
- New terms with lexical ambiguity identified:
 - Consistent, Uniform, Variable

Example Codes for Range

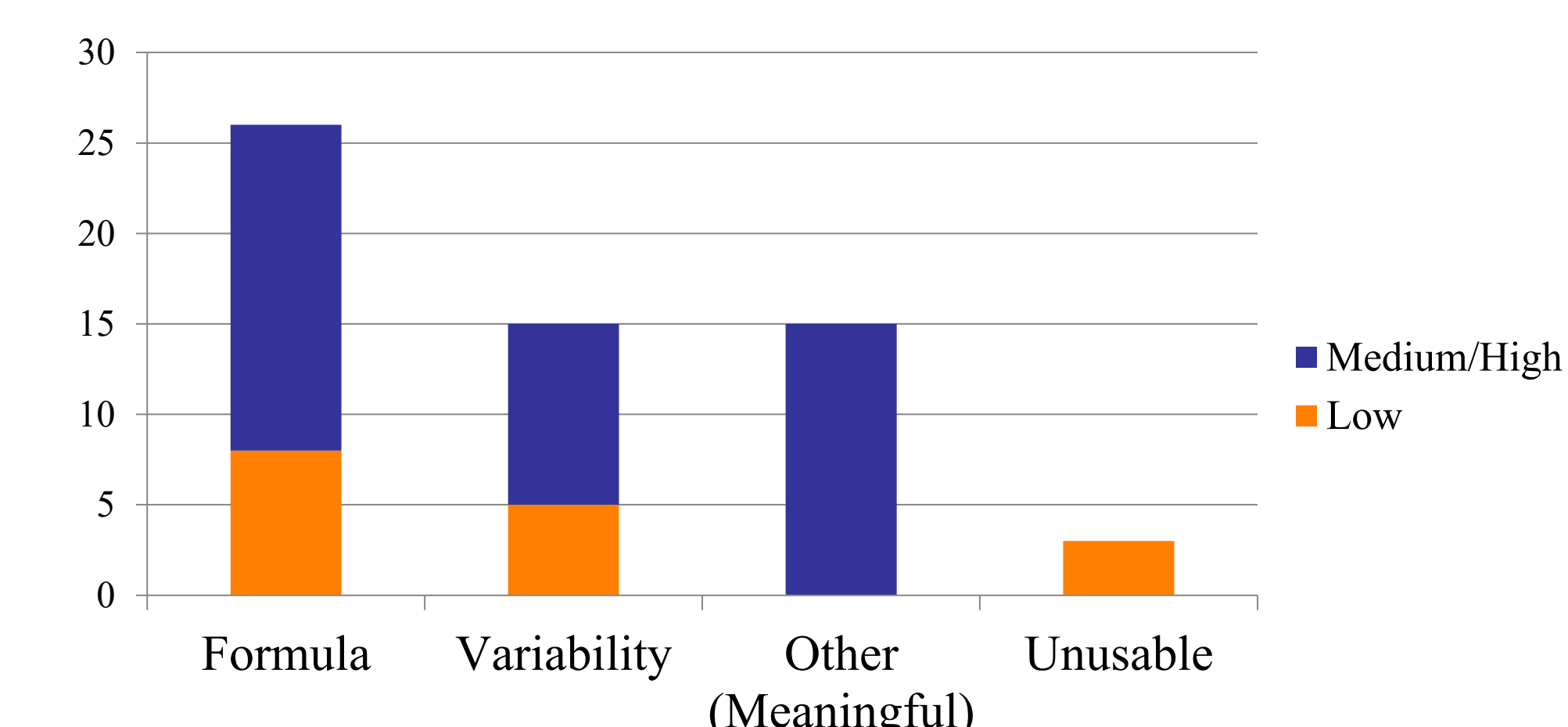
- The smallest number subtracted from the largest number.* [Formula, High]
- The range was 7* [Formula, Low]
- How spread across the data is.* [Variability, Medium]
- The range of school grades is 0-100.* [Variability, Medium]
- The range of data was large.* [Variability, Low]
- The domain and range is (4,3).* [Other (Meaningful), Medium]
- I'm gonna go to the golf range next Saturday.* [Other (Meaningful), Medium]
- a cute ranch* [Unusable, Low]

Distribution of Codes for Range

Before Instruction



After Instruction



Note: Summarized results (for both words) combine codes for definitions and sentences, and some responses included two definitions/sentences that were independently coded.