OPEN EDUCATIONAL RESOURCE

Nominal, Ordinal, Interval and Ratio: Scales of Measurement with Examples

Derek Jansen, MBA
Grad Coach

Reviewed by Eunice Rautenbach, DTech · November 2026

Abstract

Every variable in a quantitative study sits at a level of measurement, and that level determines which statistics and tests can meaningfully be applied to it. A test run on a variable at the wrong level produces output that appears sound but supports little. This article sets out the four levels of measurement described by Stevens (1946): nominal, ordinal, interval and ratio. It explains the property each level adds to the one below it. Each level is illustrated with everyday examples, and extracts from three doctoral studies show how researchers state a variable's level and tie it to the chosen test. The article then offers a procedure of up to three questions for identifying a variable's level, classifies the cases that are hardest to judge, such as age, income, IQ and Likert scales, and summarizes the debate over treating multi-item Likert scores as interval data. It closes with the tests each level typically supports and the implications for questionnaire design. It is intended for students and early-career researchers planning a quantitative analysis, and for the educators who teach research methods and statistics.

Keywords: nominal ordinal interval ratio; scales of measurement; levels of measurement; nominal vs ordinal; Likert scales; statistical tests; quantitative research; open educational resource

Suggested citation: Jansen, D. (2026). *Nominal, ordinal, interval and ratio: Scales of measurement with examples*. Grad Coach. https://doi.org/10.5281/zenodo.24197963

About this article: This is the formal edition of an article first published at gradcoach.com/nominal-ordinal-interval-ratio/, where the figures can also be downloaded for reuse. Further open educational resources on measurement and analysis are listed at gradcoach.com/open-educational-resources/.

© 2026 Grad Coach. Free to share and adapt under CC BY 4.0 with credit to Grad Coach and a link to gradcoach.com/nominal-ordinal-interval-ratio/. Quoted thesis extracts remain their authors'; each thesis keeps its own license.

1. Introduction

The level of measurement of a variable describes how much information its values carry, from simple labels to true measurements. It matters because every statistical procedure assumes a level. A t-test (a test that compares the means of two groups) on a variable that is in fact ordinal, or a mean calculated from categories that merely look like numbers, yields output that appears correct but may mislead. Examiners and reviewers commonly look for this kind of mismatch in the methods section of a thesis or research report, where each variable's level and the suitability of the analysis for it are expected to be stated.

This article is written for students and researchers preparing that statement and for those who teach it. Section 2 introduces the four levels and the cumulative relationship between them. Sections 3 to 6 describe nominal, ordinal, interval and ratio data in turn. Section 7 sets out a procedure for identifying a variable's level and addresses the cases that cause most disagreement, including Likert scales. Section 8 describes how the level constrains the choice of statistical test and the design of data collection. Throughout, extracts from three doctoral studies show how researchers have stated their variables' levels and chosen their tests.

2. The Four Levels of Measurement

Every variable collected in a quantitative study can usually be placed at one of four levels of measurement, and that level decides what can be done with it in the analysis. The classification comes from the psychologist S. S. Stevens, who set out nominal, ordinal, interval and ratio scales in 1946 and argued that each permits a different set of statistics (Stevens, 1946).

The levels are **cumulative**. Stevens (1946) noted that the operations needed to create each type of scale are cumulative: each level keeps every property of the level below it and adds one more (Table 1).

Table 1. The four levels of measurement, the property each adds, and examples.

Level	What it adds	Examples
Nominal	Categories, with no order	Blood type, native language, yes/no
Ordinal	An order or rank	Education level, satisfaction (low/medium/high)
Interval	Equal gaps between values	Temperature in °C, year of birth
Ratio	A true zero, meaning none of the thing	Age in years, income in dollars, reaction time

Nominal and ordinal data are categorical: the values are groups, even when they are coded as numbers (1 for yes, 0 for no). Interval and ratio data are numerical: the numbers themselves carry the information. This is the first distinction to establish, because it decides whether a study is **counting things or measuring them**.

The more properties a variable has, the more can be done with it. Figure 1 summarizes what each level has, permits and typically supports; it is read column by column, and each tick marks something that level has, permits or allows to be calculated. The "multiply and divide" row refers to ra-

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

tios of the values themselves, which interval data do not support; ratios of differences between interval values remain meaningful (Section 5).

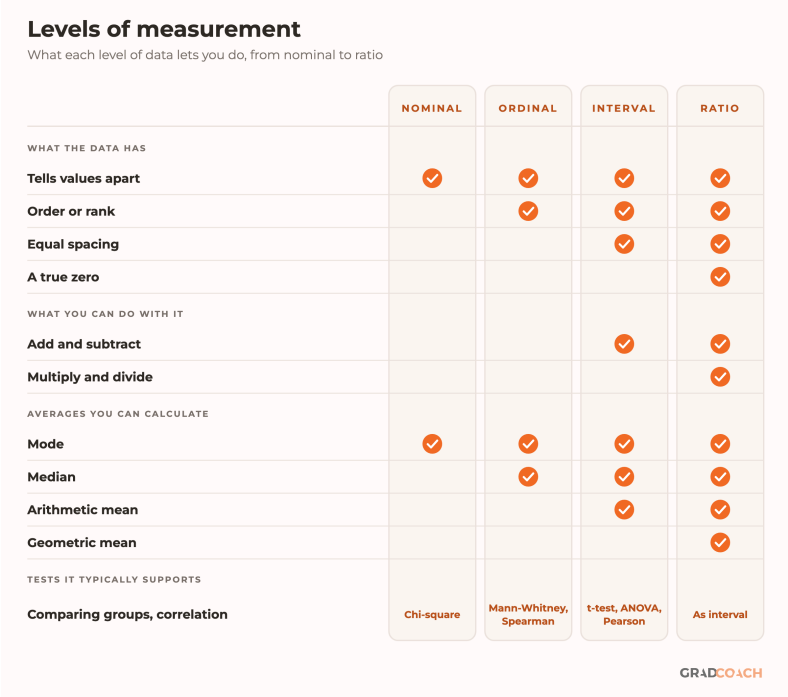

Figure 1. What each level of measurement has, permits and typically supports.

A common error is to analyze a variable as if it had a tick it lacks, most often by calculating a mean for a variable that is in fact ordinal.

3. Nominal Data

Nominal data place cases into **categories that have no natural order**. No category is higher, better or larger than another; the categories are simply different. Examples include blood type, eye color and native language; brand of phone or preferred political candidate; and any yes/no answer, such as "Have you completed a health care proxy form?"

A yes/no variable is the simplest case: two categories, no order. The UCLA statistical consulting group defines a categorical variable as one with two or more categories but no intrinsic ordering, and gives a binary yes/no question as an example (UCLA Statistical Methods and Data Analytics, n.d.-b). Coding "yes" as 1 and "no" as 0 does not change this, because the numbers serve only as labels.

With nominal data, the available operation is **counting**. A study can report how many cases fall in each category, identify the most common category (the mode) and test whether two categorical variables are related, usually with a chi-square test, which compares the counts in each combination of categories with the counts expected if the variables were unrelated (UCLA Statistical Methods and Data Analytics, n.d.-a). A mean of the category codes has no meaning, because the codes are only labels, not quantities. The one exception is a yes/no variable coded 0 and 1, whose mean is simply the proportion answering yes.

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

The following extract shows this reasoning in a nursing doctorate that tested whether a nurse-led intervention raised the share of hospital patients who completed a health care proxy (HCP) form, and then compared completion across patient groups:

> Additionally, Chi-Square test was performed to examine if statistically significant differences existed in the completion of HCP forms for gender, preferred language (English speaking v. non-English speaking), and age (≤64 years old v. ≥65 years old). Chi-square was chosen because the proxy data is nominal.
>
> Baldassari (2026, p. 16)

Two features of the extract are worth noting. First, the author **names the level and ties the test to it** in one short sentence, which is all a reader needs. A methods section should do the same for each variable: name its level, then the test that level allows. Second, age, normally ratio data, has been collapsed into two groups, which turns it into an ordered category. Data can always be moved down a level in this way, but never back up.

4. Ordinal Data

Ordinal data are also categorical, but the **categories have a meaningful order**. What they lack is equal spacing: one option is known to rank above another, but not by how much. Examples include education level (high school, bachelor's, master's, doctorate); income bands (under $30,000, $30,000 to $60,000, over $60,000); a single Likert item (strongly disagree, disagree, neutral, agree, strongly agree); and rankings, such as finishing positions in a race.

The gap between "agree" and "strongly agree" is not necessarily the same as the gap between "neutral" and "agree". Ordinal data therefore support the **median**, the middle value once cases are ranked, and **percentiles**, the values below which a stated share of cases falls (Stevens, 1946). They also support rank-based tests, which compare cases by their positions in a ranking, such as the Mann-Whitney *U* test for two groups, the Kruskal-Wallis test for three or more, and Spearman's correlation (UCLA Statistical Methods and Data Analytics, n.d.-a). Stevens (1946) observed that means and standard deviations (a measure of how far values typically lie from the mean) calculated on an ordinal scale "are in error to the extent that the successive intervals on the scale are unequal in size" (p. 679).

A public health doctorate from New Zealand, which compared Māori with favorable and adverse health outcomes using the Mann-Whitney *U* test, gives three reasons for that choice. This non-parametric test works from the ranks of the values; a parametric test, such as the t-test, works from the values themselves through the mean and standard deviation. The first reason is the instructive one, because among the variables it calls ordinal is NZDep, an area deprivation index reported in deciles, ten ranked groups numbered 1 to 10:

> This non-parametric, rank-based test was selected for three reasons. First, many of the variables of interest were ordinal (e.g., NZDep deciles, qualification ladders, cultural-identity scores), making parametric tests such as the t-test inappropriate (Hollander et al., 2013).
>
> Munn (2026, p. 125)

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

Because its groups are numbered, NZDep looks numerical. Later in the thesis, Munn describes it as "a ten-category ordinal scale rather than an interval metric" (Munn, 2026, p. 132). A number on a scale does not make a variable interval. When a variable looks numerical, the question to ask is **whether the gaps between its values are really equal**. If that cannot be shown, the variable should be treated as ordinal.

5. Interval Data

Interval data are numerical, ordered and **evenly spaced**, so the difference between two values means the same thing anywhere on the scale. What interval data lack is a true zero. As Stevens (1946) put it, "the zero point on an interval scale is a matter of convention or convenience" (p. 679). Examples include temperature in degrees Celsius or Fahrenheit; calendar years, such as year of birth; and standardized test scores, such as the SAT or GRE, which are usually treated as interval.

Zero degrees Celsius is simply the freezing point of water. For that reason, **20 °C is not twice as hot as 10 °C**, even though 20 is twice 10. Calendar years behave in the same way: the starting point of a calendar is a convention, so it is meaningless to say that one date is "twice" another. Stevens (1946) gives both temperature scales and calendar dates as examples of interval scales on which such statements are meaningless.

Because interval data can be added and subtracted, the **mean and standard deviation** are meaningful (Stevens, 1946), and so are the tests built on them: *t*-tests; analysis of variance (ANOVA), which compares the means of three or more groups; and Pearson's correlation, which measures how closely two variables move together in a straight line, provided each test's other assumptions hold. Ratios of interval values are not meaningful, so a claim that a test score of 80 shows twice the knowledge of a score of 40 is not supported. Ratios of differences are meaningful: a rise from 40 to 80 is twice a rise from 40 to 60.

6. Ratio Data

Ratio data have every property of interval data, plus a **true zero, at which there is none of the quantity** being measured. A true zero is what permits multiplication and division: 20 minutes is twice as long as 10 minutes. Examples include age in years, height and weight; income in dollars and number of children; reaction time, duration and distance; and temperature in kelvin.

Kelvin is a useful contrast with Celsius. Zero kelvin is absolute zero, the coldest possible temperature (Eckel, 2024), so 200 K is twice 100 K. Stevens (1946, p. 679) suggested that temperature on the absolute scale is probably a ratio scale. Absolute zero cannot be reached in practice, but this does not affect the classification: what matters is **what zero means**, not whether any case takes the value zero.

Ratio data support every statistic the lower levels support, plus a few that require a true zero, such as the geometric mean, an average found by multiplying values rather than adding them, which requires every value to be above zero, and the coefficient of variation, the standard deviation expressed as a proportion of the mean (Stevens, 1946, lists the latter). In practice, interval and ratio

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

data are analyzed with the same tests, and the statistical package SPSS places them under a single setting (Section 7.2). The difference appears in how results are described: a report that one group scored “twice as high”, or a percentage change, requires **a ratio variable**.

7. Identifying the Level of a Variable

Most variables can be classified with two or three questions. Figure 2 sets these out as a flowchart, to be followed from the top for each variable in a study.

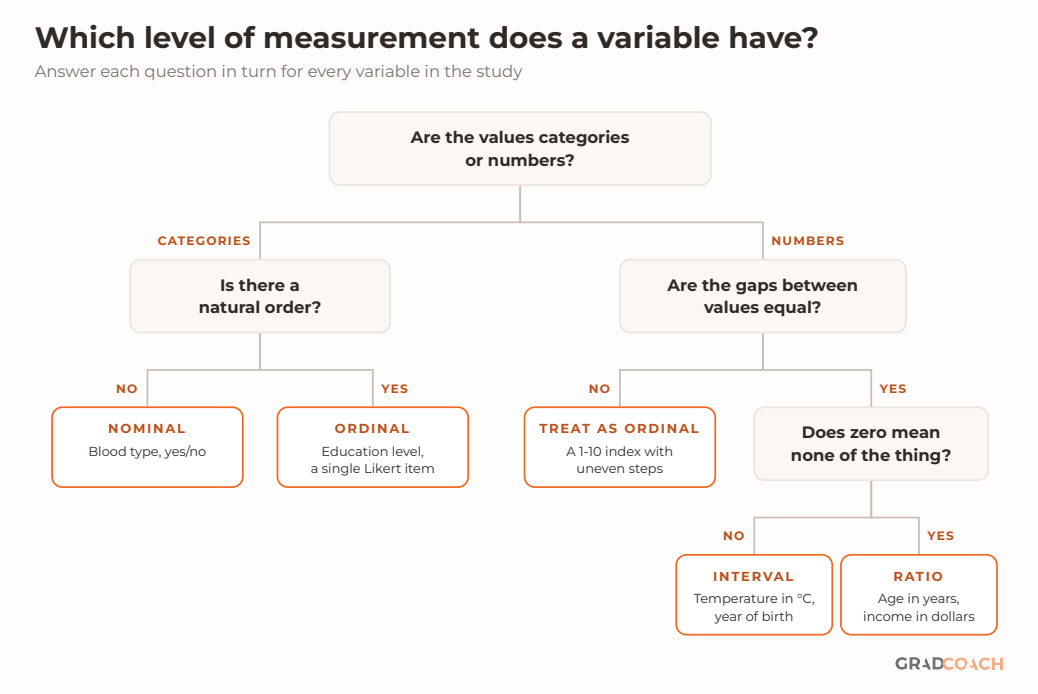

Figure 2. Identifying a variable’s level of measurement in up to three questions.

The question most often skipped is the one about equal gaps. Numbers with unequal gaps **belong on the ordinal branch**, however much they resemble a measurement. Every variable should pass through the flowchart before any test is chosen.

Some variables are harder to classify, and these are the ones most often queried by examiners and reviewers. Table 2 gives the usual level for each and the reason.

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

Table 2. Variables that are hard to classify, with their usual level of measurement.

Variable	Usual level	Reason
Age in years	Ratio	Zero years means no time has passed
Age in bands (18–24, 25–34 …)	Ordinal	Ordered groups, but the bands hide the gaps
Year of birth	Interval	Equal gaps, but the calendar's starting point is a convention
Income in dollars	Ratio	$0 means no income
Income bands	Ordinal	Ordered groups; the bands hide the exact amounts
A single Likert item	Ordinal	Ordered answers, gaps not known to be equal
A score summed across several Likert items	Debated	Often treated as interval (Section 7.1)
IQ score	Debated	Often treated as interval; arguably ordinal
Letter grades (A to F)	Ordinal	Ordered, but an A and a B are not a set distance apart
Percentage exam score	Debated	Equal steps in marks, but 0% is not zero knowledge
Gender, ethnicity	Nominal	Categories with no order

IQ scores are commonly treated as interval. Stevens (1946) himself took a more cautious view, writing that "intelligence, for example, is usefully assessed on ordinal scales which try to approximate interval scales" (p. 679). A study that relies on IQ scores should acknowledge this debate in a sentence. The Likert rows require a fuller treatment, given in Section 7.1, and Section 7.2 describes a software default that can misclassify a variable before any analysis begins.

7.1 The Likert debate

A single Likert item is ordinal, and few dispute this. The disagreement concerns what happens when several items are **summed or averaged into a scale score**.

One position, argued by Jamieson (2004) in *Medical Education*, holds that Likert scales fall within the ordinal level of measurement because the intervals between response categories cannot be presumed equal, and that they should therefore be analyzed with ordinal methods. The opposing position holds that a scale built from several items behaves well enough to be treated as interval. Carifio and Perla (2007) write that "many studies have shown that Likert Scales (as opposed to single Likert response format items) produce interval data" (p. 106), and Norman (2010) cites studies dating back to the 1930s showing that parametric tests give dependable results even when assumptions of this kind, such as equal intervals, are not fully met.

Jamieson (2004) herself reports the observation that it has become common practice to assume Likert-type categories constitute interval-level measurement (Blaikie, 2003, as cited in Jamieson, 2004), and many doctoral studies do so. A doctoral dissertation in agricultural education from

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

Purdue University illustrates a practice worth adopting: it built its scales from five-point Likert items and then set out, in a single table, each research question's variables, their levels and the planned analysis. Table 3 reproduces three of its rows in shortened form.

Table 3. Variables, levels and analyses for three research questions (Huang, 2019).

Research question (shortened)	Variables and their level	Data analysis
RQ3: how perceptions, belonging and self-efficacy relate	Perceptions of teaching practices, sense of belonging, academic self-efficacy: **interval**	Mean; standard deviation; correlations; path coefficient
RQ4: differences by race, classification, gender, college and institution type	The five grouping variables: **nominal**. Perceptions of culturally responsive teaching: **interval**	Mean; standard deviation; ANOVA (3+ groups); independent samples t-test (2 groups)
RQ5: other factors that shape perceptions	Open-ended answers: **nominal**	Codes and themes

Note. Adapted from Huang (2019, Table 1, p. 80). Columns combined and wording shortened; the original lists independent and dependent variables separately and leaves the RQ5 variable cells blank.

A table of this kind gives a reader one place to check every variable's level against its test. The scale scores in it, however, are treated as interval without a stated rationale. The thesis did check the scales' structure by factor analysis, a method that tests whether items group together as intended, but that check does not address the level. The practice is common and usually accepted, but it leaves the choice undefended. Where a scale has several items and a published validation, the score can reasonably be **treated as interval, with the decision stated and supported by a citation**. Where the measure is a single item, ordinal methods are the safer choice.

7.2 Levels of measurement in SPSS

SPSS records each variable's level in its Measure column, which offers three settings: **Scale**, **Ordinal** and **Nominal** (Kent State University Libraries, n.d.). Interval and ratio data share the Scale setting; there is no separate option for each.

By default, SPSS sets variables with numeric values to Scale (Kent State University Libraries, n.d.), so **a nominal variable coded 1, 2 and 3 may be treated as Scale** unless the researcher changes the setting. The Measure column is therefore the first thing to check in a new data file.

8. Implications for Analysis and Design

A variable's level of measurement narrows the choice of statistical test. Most studies need to compare groups or test a relationship. For these two tasks, the usual guidance by the level of the outcome, the variable being compared or explained, is as follows (UCLA Statistical Methods and Data Analytics, n.d.-a):

1. **Nominal outcome:** a chi-square test.
2. **Ordinal outcome:** the Mann-Whitney *U* test, or the Kruskal-Wallis test for three or more groups; Spearman's correlation for relationships.

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

3. **Interval or ratio outcome:** a t-test, or ANOVA for three or more groups; Pearson's correlation for relationships, provided the tests' assumptions hold.

The source of this mapping describes it as "general guidelines and should not be construed as hard and fast rules" (UCLA Statistical Methods and Data Analytics, n.d.-a, para. 1), and the caveat deserves weight. In a well-known critique of Stevens's typology, Velleman and Wilkinson (1993) argued that the typology is misleading as a guide to choosing statistics, and that it fails to answer criticisms raised when it was first proposed. The levels are therefore best used as a starting point, with **any choice that departs from them justified** in the methods section.

The level of measurement also matters before any data are collected. A study that will require a t-test should be **designed to produce numerical data**: asking for age in years rather than in bands, and income as an amount rather than a range, where this is feasible. Ratio data can always be grouped later; grouped data cannot be ungrouped.

9. Conclusion

Nominal, ordinal, interval and ratio levels are cumulative: each keeps the properties of the level below it and adds one. Nominal and ordinal data are categories that are counted or ranked; interval and ratio data are numbers that are measured. Interval and ratio data differ in the meaning of zero: only on a ratio scale does zero mean none of the quantity, so that one value can be described as a multiple of another. A single Likert item is ordinal, while the status of summed scale scores is debated and any decision to treat them as interval should be justified. Data are best collected at the highest level feasible, and a methods section should state each variable's level alongside the test it supports.

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com

References

Baldassari, J. (2026). *Improving health care proxy procurement rates for hospitalized adults: A nurse led intervention* [Doctor of Nursing Practice project, University of Massachusetts Amherst]. https://doi.org/10.7275/10651

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 3(3), 106–116. https://doi.org/10.3844/jssp.2007.106.116

Eckel, S. (2024, June 26). How low can temperature go? Lord Kelvin and the science of absolute zero. *Taking Measure*. National Institute of Standards and Technology. https://www.nist.gov/blogs/taking-measure/how-low-can-temperature-go-lord-kelvin-and-science-absolute-zero

Huang, Y. (2019). *Undergraduate students' perceptions of culturally responsive teaching and their sense of belonging and academic self-efficacy in higher education* [Doctoral dissertation, Purdue University]. Purdue University Graduate School. https://doi.org/10.25394/PGS.9116885.v1

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Kent State University Libraries. (n.d.). *SPSS tutorials: Defining variables*. Retrieved November 11, 2026, from https://libguides.library.kent.edu/SPSS/DefineVariables

Munn, S. (2026). *Understanding survival and suffering: An individual-level joint-burden approach for measuring health loss among Māori and non-Māori: Illustrated using a national retrospective cohort study of type 2 diabetes in Aotearoa New Zealand* [Doctoral dissertation, Te Herenga Waka-Victoria University of Wellington]. Open Access Te Herenga Waka-Victoria University of Wellington. https://doi.org/10.26686/tjqy-59wz

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677–680. https://doi.org/10.1126/science.103.2684.677

UCLA Statistical Methods and Data Analytics. (n.d.-a). *Choosing the correct statistical test in SAS, Stata, SPSS and R*. Retrieved November 11, 2026, from https://stats.oarc.ucla.edu/other/mult-pkg/whatstat/

UCLA Statistical Methods and Data Analytics. (n.d.-b). *What is the difference between categorical, ordinal and interval variables?* Retrieved November 11, 2026, from https://stats.oarc.ucla.edu/other/mult-pkg/whatstat/what-is-the-difference-between-categorical-ordinal-and-interval-variables/

Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47(1), 65–72. https://doi.org/10.1080/00031305.1993.10475938

© 2026 Grad Coach · CC BY 4.0 · gradcoach.com