Demystifying Surveys

A Guide to Using MR Measurement Scales

By Yusof Ahmad, Head of Methodology and Director, Insights

PHRONESISPARTNERS
simplifygrowth

Foreword

Market research is pivotal in guiding business strategies through data-driven decisions, but the importance of accurate measurement scales is often underestimated. In a world flooded with data, the choice of the right measurement scale can mean the difference between insightful analysis and misguided conclusions.

We've created this guide to highlight the vital role measurement scales play in deciphering the complex world of surveys, providing researchers, analysts, and decision-makers with a comprehensive roadmap to navigate measurement scales effectively.

We cover not only the range of key measurement tools available, but also their distinctions and practical applications. By shedding light on their distinct characteristics and optimal usage scenarios, this guide is designed to help researchers make informed choices about their survey models and methodologies to help them achieve their business objectives. It explores important considerations such as the balance, nature, and types of scales and their implications for analysis.

As the first paper in a series on Research Methodology, it establishes a framework from which to explore advanced analytical techniques for measuring and assessing market research, and gain meaningful insights from survey data. Subscribe to get notified when future articles in this series are published.

Binayak Choudhury
Managing Partner
PHRONESIS PARTNERS

More about the series:

- This series provides a roadmap to help professionals navigate the intricacies of market research for a rapidly changing landscape where success relies on accurate data-driven insights.
- We guide you through the practical aspects of market research, going beyond theories to address real-world considerations, sharing the key steps to creating impactful market research.
- Each guide in the series contains relatable examples designed to cement your understanding of how to achieve the results you need to drive your business goals.

CONTENTS

1 Understanding measurement scales in Market Research

In market research, the term "scale" has two meanings. One refers to multi-item measures (i.e. SERVPERF or SERVQUAL), risk propensity scales, or brand personality scales. These measure intangible concepts using multiple items, each with its own scale. This guide, however, looks at the other types of "measurement scales" - those that measure perceptions, attitudes, or behaviors for a single variable. Among these, we'll primarily cover rating scales since they're most popular.

Continuous Rating scales: Imagine a straight line where participants indicate their answers. They can choose a marked point or anywhere in between, with the system calculating a number based on their selection within the given range. Often, this straight line might only have labels at its endpoints. For example:

QA10. How likely are you to start working with NoSQL databases in the next 3 months? Slide the point on the line below. Closer to 100 means more likely.

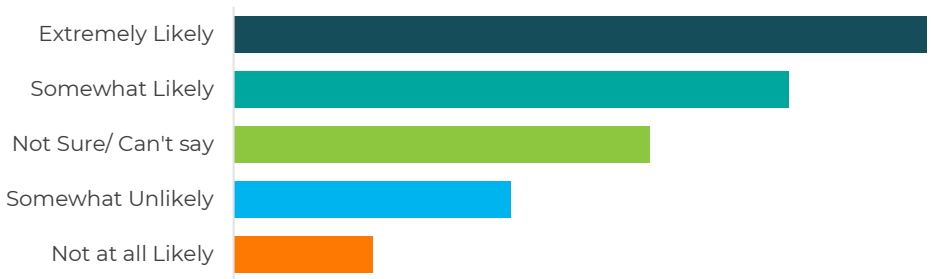
While these scales can capture decimal numbers, it's more common to use a broader range, like in the previous example. The many scale points allow it to function as a continuous scale for all purposes.

1. **Itemized Rating scales:** Participants choose from a set of labeled points arranged in order. As you move from left to right on the scale, the value either increases or decreases. This is the most widely used scale in market research.

Now, let's look at some popular itemized rating scales:

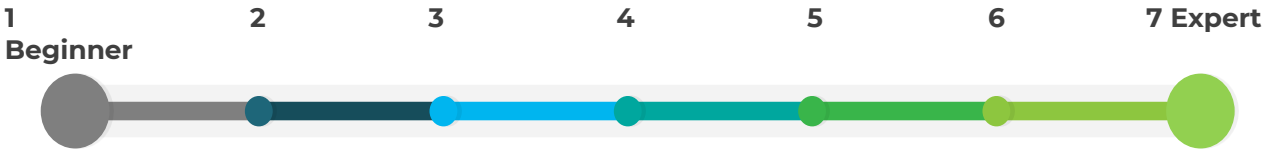
Likert-type scale: We've all seen these. They're labelled at each point, with an equal balance of positive and negative items around a neutral zero point. The labels are called verbal anchors. For example, a scale could range from "Not knowledgeable at all " to "Highly knowledgeable." Below is an example of a Likert-type scale used for the same question as in the previous example:

QA10. How likely are you to start working with NoSQL databases in the next 3 months? Please select an option below:



2. **Semantic Differential Scales:** These use polar opposite verbal anchors, allowing respondents to indicate their attitude or evaluation. An example of this might be asking respondents to evaluate a product idea using a scale from "Novice" to "Expert".

QA11. How would you rate your knowledge of NoSQL database programs? Please answer using the scale below, where higher ratings mean more knowledge.



3. **Stapel's Scale:** This unique unipolar even-point non-verbal rating scale ranges from -5 to +5 and has no neutral zero point. We can use a 10-point Stapel's scale for our expertise question in the previous example:

QA11. How would you rate your knowledge of NoSQL database programs? Please select one of the numeric options below.

We can also use a 6- or 8-point scale in the above example. The rule is that the number of positive and negative points should always be equal.

2 The nature of different rating scales

When respondents use rating scales to respond to questions about their attitudes or beliefs, they consider both the extreme points of the scale (position effect) and the verbal anchors (in Likert and semantic differential scales).

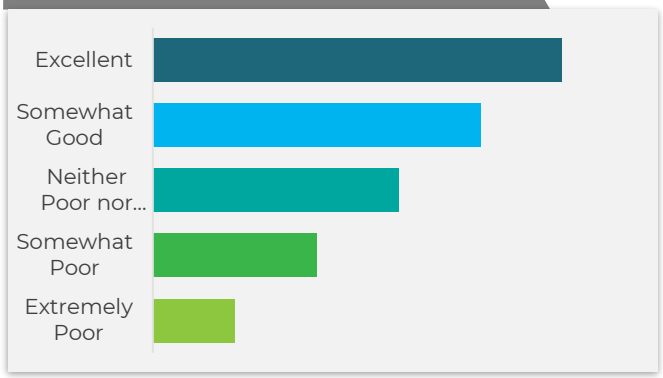
Let's now look at rating scales in terms of the nature of the scale itself:

Balanced vs Unbalanced Scales

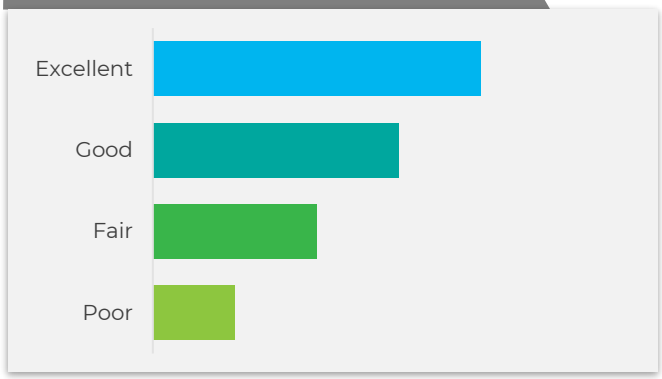
Balanced scales have a neutral midpoint with equal positive and negative points on either side. Unbalanced scales lean more towards positive or negative points. We often see scales tilted positively, like "Poor", "Fair", "Good", "Very Good", and "Excellent" which has more positive than negative points. Balanced scales are usually preferred as unbalanced ones might hinder honest responses. Here are some examples to highlight the differences:

QA12. How would you rate the performance of ChatGPT-3.5 in solving roadblocks you face while coding with Python? Please select an option below:

Balanced



Unbalanced



Ordinal vs Interval Scales

Most scales we've discussed are ordinal, showing higher or lower levels of a trait without specifying the exact difference. Yet, these ordinal scales are often treated as interval scales, implying that the numeric differences have meaning. So, a 1-point difference anywhere on the scale is assumed to represent the same change in beliefs or attitudes. For instance, the gap between 7 and 6 is seen as equivalent to the gap between 5 and 4 on a 7-point scale.

3 Analysis and interpretation of results from rating scales

Let's explore how to analyze responses to rating scale questions. Since scales are typically ordinal but treated as interval, these basic analysis methods are commonly used to interpret results:

1. When the scale is treated as an ordinal scale:

a) Summarize responses by the percentage at each scale point (e.g., from our example of the rating of ChatGPT-3.5 in helping resolve issues with Python code):

Rating	Frequency	Percentage
1 = Extremely Poor	30	15.00 %
2 = Somewhat Poor	35	17.50 %
3 = Neither Poor nor Good	40	20.00%
4 = Somewhat Good	55	27.50%
5 = Excellent	40	20.00%
Total	**200**	**100%**

b) Summarize responses by the percentage at the top-n or bottom-n boxes. Using the above table:

Category	Frequency	Percentage
Top-2 Box (4+5)	95	47.50 %
Mid Box(3)	40	20.00%
Bottom-2 Box (1+2)	65	32.50%
Total	**200**	**100%**

As can be seen, this is more convenient as it summarizes the first table and clearly tells us that nearly half of all respondents feel that ChatGPT-3.5 is helping them deal with roadblocks with their Python scripts.

c) Compare percentages at the top-n and bottom-n boxes as summarized above, between two or more groups (e.g., App Developers and Data Analysts). Let us look at some hypothetical data again:

Category	App Developers	Column Percentage	Data Analysts	Column Percentage
Top-2 Box (4+5)	95	47.50 %	52	52.00%
Mid Box (3)	40	20.00%	8	8.00%
Bottom-2 Box (1+2)	65	32.50%	40	40.00%
Total	**200**	**100%**	**200**	**100%**

If the data were purely categorical (i.e. Vote Republican, Democrat, Greens, etc.), we'd use a chi-square test to see if there's a relationship between categories and the groups. However, in our scenario, data is ordered (Top-2 box is more positive than mid-box, which is better than bottom-2 box). For example, data analysts show higher scores in both top-2 and bottom-2 boxes.

To determine if data analysts view ChatGPT-3.5 more favorably than App Developers, we use the chi-square trend test. These tests give a p-value, and a value below 0.05 suggests a notable difference between groups. If we hypothesize one group's percentage is more or less than another, the test adjusts accordingly. With more than 2 groups, additional modifications to the chi-square test are needed. These tests and tweaks can be done using statistical software like SPSS, STATA, or data analytic tools in R and Python.

d) Compare whether two groups of respondents have similar or dissimilar rankings/preference orders. If the scale is ordinal, we must use measures of association that do not assume an underlying distribution in the data. Distribution-free tests of association like Spearman's Rank correlation and Kendall's Tau are used.

Spearman's Rank Correlation and Kendall's Tau use the ranks of responses rather than the raw data. This allows researchers to determine the degree to which the order of rankings from two different groups align, offering insights into whether there is a relationship between the two groups' attitudes, behaviors, or perceptions.

2. When the scale is treated as an interval scale:

a) Summarize responses by statistical descriptive measures like the mean and standard deviation.

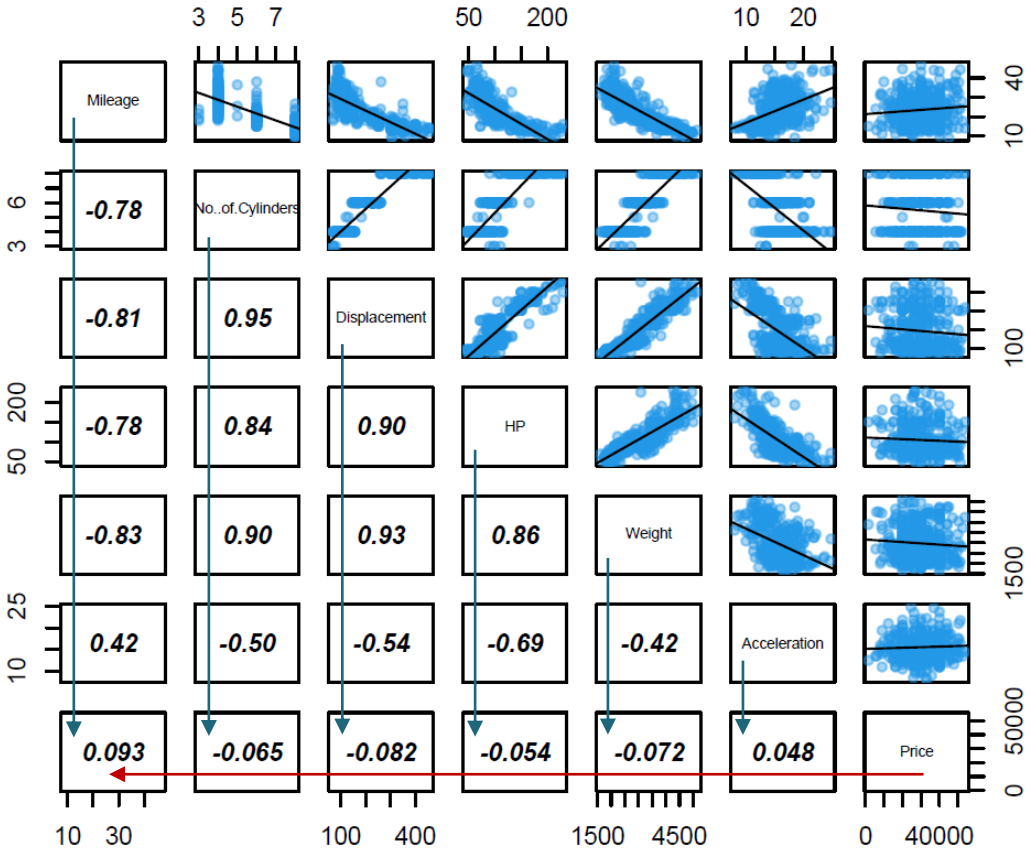
b) Use the t-test of differences in means or ANOVA to compare between groups (if there are more than two groups).

c) Use correlations to determine the relationship between two variables. The Pearson Correlation is commonly used to see if two variables are linearly related. It ranges from -1 to +1. Positive values indicate both variables increase together, while negative values mean one variable rises as the other falls. A value of +1 or -1 indicates a perfect relationship, like when one variable's values are consistently double the other's.

T-tests and Analysis of Variance (ANOVA) help researchers determine whether there are significant differences in average responses across groups (for instance, whether one product is rated significantly higher on average than another by consumers), allowing for evidence-based decision-making.

4 Pearson correlations

The following diagram shows the correlations between pairs of variables of a dataset of car features from a publicly available dataset. We will read off the value of the correlations from the intersections of the rows and columns in the bottom panel of the matrix and look at the bivariate plots from the intersection of the rows and columns in the upper panel of the matrix. The variables are named along the diagonal.



The displayed diagram, or correlation matrix plot, consists of an upper and lower section. The lower section displays numeric correlation values, while the upper shows bivariate plots. Variable names run diagonally.

To find the correlation between two variables, locate the value where one variable's column intersects with the other's row in the lower section. For the plot, do the same in the upper section. For instance, the correlation between price and HP is a slight -0.054, shown by the intersection of the orange lines. In the upper section, the HP vs. Price plot is roughly shapeless, lacking a clear trend. The points spread evenly around a horizontal line, the regression line, revealing no obvious trend.

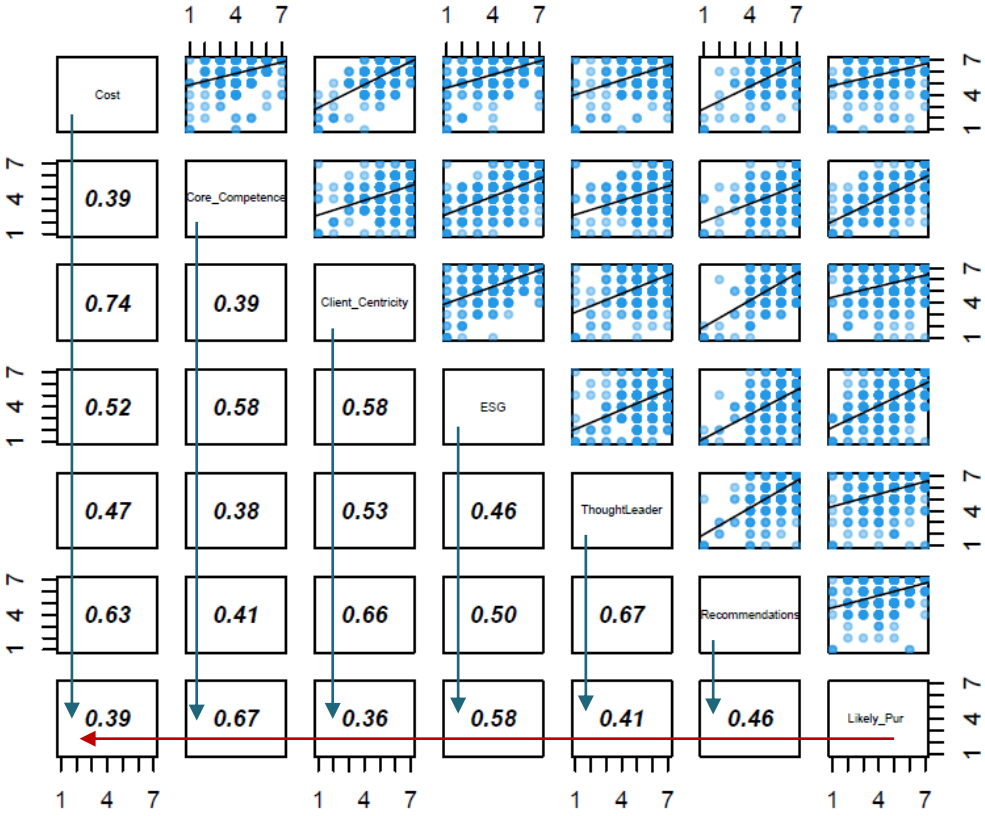
In contrast, "Horsepower" vs. "Displacement" has a strong correlation of 0.90. The related scatter plot reveals points closely following an upward-trending line. This trend is mirrored in "Horsepower" vs. "Weight" and "Weight" vs. "Displacement".

However, plots against "No. of Cylinders" differ. Since the number of cylinders takes only specific integer values (2 to 8 in our data), plots aren't continuous. We can still deduce correlation assuming continuity. Despite being discrete, we can fit a straight line to represent an average linear relationship. For “No. of Cylinders” vs. “Displacement”, “HP”, and “Weight”, we observe an upward trend with corresponding positive, high correlation values.

5 Correlation analysis when using Ordinal MR scales

Correlation analysis is frequently used to identify relationships between variables, but it can be misleading and lead to wrong conclusions. Especially with Likert-type scales or other discrete measures, it's essential to recognize their limits and consider better-suited methods like polyserial or polychoric correlations.

We'll explore these methods in more detail later. For now, bivariate correlations are sometimes applied in key driver analyses. Here, a dependent variable like "willingness to purchase" (measured continuously) is compared with potential driver variables. The correlation strength is seen as an indicator of each driver's importance, though this approach has its flaws.

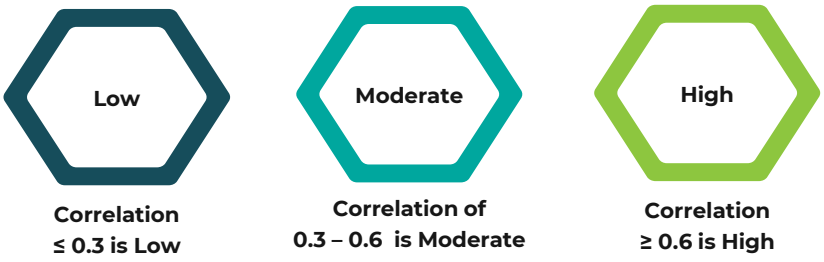


Let us say all decision-makers evaluated five firms using 7-point scales, and we have their ratings on the following attributes of the supplier firms:

- Cost
- Core Competence
- Client Centricity:
- ESG
- Thought Leadership
- Recommendations

We also have data on their "willingness to purchase" (Likely_Pur) from each of those firms. Examining the correlation matrix, we see scatterplots showing discrete variables clustered around distinct values on the x- and y-axes. There's no linear relationship between any two variables. However, if we assume an underlying continuous distribution, we can fit a straight line to each plot's point distribution. The steepness of these lines indicates the correlation strength – steeper lines signify higher correlations. In this case, all lines slope positively.

Observing the horizontal blue arrow in the “scatter plot matrix”, most of the correlations do not seem strong; most are below 0.5. Market researchers often apply specific thresholds, such as:



At Phronesis Partners, we believe there's a better approach to examining key drivers than just observing bivariate correlations. These correlations aren't clear indicators of association, especially when multiple drivers correlate with each other. For instance, the low correlation between Cost and purchase likelihood seems counterintuitive. It could be due to Cost's strong correlation with other factors that also influence Purchase Likelihood.

To truly understand the impact of Cost on Purchase Likelihood, we need to isolate its effect, ensuring other variables remain constant. This leads us to delve into Regression Analysis and Partial Correlations, which we'll discuss in a future article. Suffice it to note here that the lines that we had in our correlation matrix plot were regression lines for regression between two variables where such confounds were absent (or made absent because they had been ignored by using a two-variable model)

6 Advanced correlation methods

Pearson Correlations gauge linear associations for continuous, interval-scaled variables. In market research, scales with 7 points or more often act as interval and continuous scales. But with 3 or 5-point scales, this assumption becomes shaky. When assessing associations between two binary attributes, we face challenges. For instance, if we had asked in our IT solution example whether certain drivers were present or absent, how do we measure this association? One approach assumes binary responses stem from observations on a continuum. Using this assumption, statisticians have proposed several correlation types:

Polyserial correlations

Used when one variable is on a discrete, ordinal scale like a Likert-type scale, and the other is continuous. For instance, correlating revenues in millions USD with next year's Business Sentiment on a 5-point Linkert-type scale.

Polychoric correlations

Applied when both variables are ordinal and categorical. An example would be correlating Customer Satisfaction and Willingness to Purchase, both measured on a Likert-type scale.

Tetrachoric correlations

A special case of polychoric, used for two binary variables.

Further details on these will be covered in future guides. However, do note that these options are available with most statistical software packages or data analysis packages for R and Python.

7 Testing significance In correlations

Reporting significance for correlation coefficients is common. Unless a particular hypothesis is set, a significant correlation simply means we have enough evidence to dismiss the idea that the correlation is zero. This means the observed correlation isn't just by chance due to the specific sample we've taken. However, it's often more valuable to consider the size and direction of the correlations. Using thresholds helps in this regard. With large samples, even tiny correlations can appear statistically significant, though they might not be practically significant.

Measurement scales in market research play a crucial role in collecting reliable data. This article introduced various scales, from the basics to more advanced methods. While these tools are valuable, it's important for researchers to use them correctly to get accurate results. As we continue to develop this series of articles we will dive deeper into more advanced research methods.

About Phronesis Partners

Phronesis Partners, including our group companies, Simplify Growth and Fuld & Company, is a global research and analytics firm. We deliver high-quality, best-in-class corporate, industry and market research to many of the world's largest research and consulting firms, private equity firms, financial institutions, and publishers.

Our research is both diverse and impactful, helping clients address many of their most complex problems. Be it understanding performance improvement areas for a Bauxite mine in Africa, or building a complex financial model for a Silicon Valley unicorn, or developing a comprehensive Brand index for the world's top brands, our clients rely on us for some of the most complicated research questions across industries.

Contact Us

Yusof Ahmad
Head of Methodology and Director – Insights
yusof.ahmad@phronesis-partners.com

Binayak Choudhury
Managing Partner
binayak.choudhury@phronesis-partners.com

phronesis-partners.com

Register for our newsletter and to receive company updates