Meta Analysis · 3
How to assess the strength of relationships

Chuck Chakrapani

The strength of relationships
In marketing research, we are often interested in finding out the strength of the relationship between two variables. To what extent does improved quality result in increased customer satisfaction? How much does lowering the price lower the perceived quality of a product? Is overall liking for a product strongly or weakly related to purchase intent? How much is advertising exposure related to product purchase?

We generally measure these variables using rating scales, which are generally presumed to be interval scales. In some instances, such as sales revenues, we measure the variables as ratio scale items. When this is the case (i.e. variables being measured as interval or ratio scales) finding the overall correlation across different studies is straight forward. It is simply the mean of (Pearson product moment) the correlations obtained in the different studies.

Let us start with an example. Four small studies attempted to identify the relationship between a person's risk aversion (as measured by a battery of questions) and the total amount of insurance bought by that person. The results are given in Exhibit 1. The correlations are not consistent. They range from a low of -0.31 to a high of +0.66. What can we say about the nature of the relationship between risk aversion and the likelihood of purchasing insurance when we combine all four studies? As mentioned earlier,

Overall correlation
= Sum (correlations)
÷ Number of studies

= (.66 + -.31 +.49 +.15) / 4 = 0.25

Similarly, to obtain the corresponding Z score, we simply obtain the mean of the individual Z scores.

Overall Z score
= Sum (Z scores) / Number of studies
= (.79 -.32 +.54 +.15) / 4 = 0.29

Which should we prefer - r or Z? While there is some evidence that Z tends to slightly overestimate the population r, the differences are negligible except in cases where the sample size is small and the population r is large. You may also want to consider weighting each correlation by the sample size of the study on which it is based.

Exhibit 1
Correlations between Risk Aversion and Total Amount of Insurance Held

Study	**N**	**r**	**Z**
1	220	.66	.79
2	250	-.31	-.32
3	180	.49	.54
4	150	.15	.15

Interpreting the correlation

Is the correlation obtained by combining the four studies large? In answering this question, it is preferable to have an understanding of the field we are working in. For instance, if we are working in the insurance field, other potential variables such as income or occupation may be only slightly correlated with total amount of insurance bought (say between .1 to .2). In this situation a correlation of .25 between risk tolerance and the total amount of insurance bought can be considered large. If, on the other hand, other variables such as the number of dependents, disposable income or age correlate strongly with amount of insurance bought (eg. .6 to .8), then a correlation of .25 can be considered rather low.

Rosenthal and Rubin provide a statistical test to assess the importance of the correlation coefficient. The procedure consists of transforming the obtained r to a c2. This is then converted to a Binomial Effect on Display (BESD). The BESD is the estimated difference in success probabilities.

Let us apply this to our example. If we were to divide Risk Tolerance scores as being above (Risk Averse) or below (Non-risk Averse) the median and the total amount of insurance bought as high or low, how much will a person's being above the median, in Risk Aversion, increase his/her probability of being in the high insurance group? This is given by

$50\% \pm (r/2)*100$

If Risk Tolerance is unrelated to the Total Insurance Amount, then 50% of the customers would be low insurers and 50% would be high insurers, irrespective of their risk tolerance. This probability would change if the two variables were correlated as happens in our example. Since the obtained correlation in our example is .25, applying the formula above, we get

$50\% \pm (.25/2)*100 = 37.5\%, 62.5\%$

This means that as the probability of one's buying increases from 37.5% to 62.5%, one moves from Non-risk Averse status to Risk Averse status (Exhibit 2). In other words, a correlation of .25 translates into an increased probability of .25 (25%) between the two groups. Since an increased probability of .25 can have major marketing implications, a correlation of this magnitude (especially when it is based on a number of studies) can be considered large. This paints a somewhat different picture compared to the standard Coefficient of Determination or r-squared interpretation-an r of .25 'explains only 6% of the variance. This is presented in Exhibit 3.

Exhibit 2
Binomial Effect Size for r=.25
Between Risk Aversion and Insurance Amount

	Amount of insurance	
	High	Low
Non-risk averse	37.5%	62.5%
Risk averse	62.5%	37.5%

Exhibit 3
Correlation coefficients and
Binomial Effect Size

Correlation	Probability *From*	Probability *To*	Increased probability	Variance explained
.10	.45	.55	10% (.10)	
.20	.40	.60	20%	1%
.25	.38	.63	25%	4%
.30	.35	.65 .70 .75	30%	6%
.40	.30	.80 .85 .88	40%	9%
.50	.25	.90 .95	50%	16%
.60	.20		60%	25%
.70	.15		70%	36%
.75	.13		75%	49%
.80	.10		80%	56%
.90	.05		90%	64%
				81%

Reconciling different statistical measures
Not all studies use the same statistical measures. This can be a problem if we do not have access to the original data from which the results we are attempting to integrate came from. Obviously, if we need to combine several studies, we need techniques that will enable us to convert them into a standard measure such as r or d. Procedures are available to do this. Exhibit 4 shows some common procedures that can be used to convert a variety of statistics such as t, F and c2 . Once we decide on a common statistic such as r or d, we need to convert the result of all studies under analysis to that statistic. As can be seen from the formulas, the calculations are fairly simple and straight foward.

Exhibit 4
How to Convert to Common Statistics

Original	*Conversion formulas* *To convert to r*	*To convert to d*
t	sq.rt. [t^2 / (t^2 + df)]	2t / sq. rt.df
F	sq.rt.[X^2/ (F+ df_{error})]*	2 sq.rt.[F/(df_{error})]*
X^2	sq. rt.(X^2 / n)**	
r		2r / sq. rt.(1 - r^2)
d	d / sq. rt. (df + 4)	

* For comparing 2 means

** For 2x2 tables

How to combine different populations
There is one other problem. Not all studies are based on the same target audience. In our example we combined different studies to understand the correlation between risk aversion and insurance buying. The basic assumption is that all studies had the same target population. However this may not always be the case.

What if we have access to a number of studies and some of the studies were conducted in Quebec and others in Ontario? It is possible that the relationships are not identical.

Let us consider an example in which a brewery conducted a series of studies to assess the correlation between alcoholic content of beer and overall liking. There were 18 studies in all, 12 of which were conducted in Ontario and 6 in Quebec. We compute the effect size and find that it is .24 for Ontario and .36 for Quebec. Now we can create a binary variable by setting Quebec to 1 and Ontario to 0. This binary variable can then be correlated with the effect size for each of the studies. If the correlation turns out to be substantial and statistically significant, we can then conclude that the relationship between alcoholic content and overall liking is higher in Quebec than in Ontario.

An even simpler approach to assessing the difference in effect sizes is to use a straight forward test of significance for effect sizes obtained in Quebec and in Ontario. Thus:

Z = (Zr1 - Zr2)
÷ [(1/(N1-3) + (1/(N2-3)]

The statistics relating to the brewery study described earlier are outlined in Exhibit 5. We can apply the formula for the meta-analysis to this data in Exhibit 5. (The average correlation of .24 and .36 is equal to a Z score of 0.25 and 0.38 respectively).

Z =
(.25 - .38) / [(1/9) + (1/3)]
= - 0.20

A Z value of -0.20 is not significant at the 95% level. Therefore we conclude that the correlation between alcoholic content and overall liking is not stronger in Quebec than in Ontario.

Exhibit 5
Mean Effect Size For
Different Regions

Region	**N**	**r**	**Z**
Ontario	12	.24	.25
Quebec	6	.36	.38

Suppose we calculate d values instead of r values. In this instance, we use d values rather than r values in our calculations. The calculating formulas themselves do not change, except for the substitution of d values for r values.

When we have a number of mediating variables
When we have a number of mediating variables (variables that differ from one study to another) influencing the outcomes, we can also use a multiple regression analysis approach through which we can assess the influence of a number of mediating variables.

Illustrating
Computer Documentation
The Art of Presenting Information Graphically on Paper

by William Horton
Published by John Wiley
New York, 1991

Why is a book like 'Illustrating Computer Documentation' being reviewed in a newsletter like Imprints? This is definitely one of those cases where one shouldn't judge the book by its cover (title?). The book has more to do with its subtitle 'The Art of Presenting Information Graphically on Paper" than with its main title.

With the advent of personal computers and graphic software, many reports these days are filled with charts and graphs. The only problem is that most of the graphs are drawn for decorative purposes with an intent to impress rather than for communication purposes with an intent to simplify. The result is that most graphs-be they in research reports or in newspapers-are colourful and pretty but shallow and confusing for those who want to understand what is being communicated .

Here are some examples of the common mistakes in presenting information in graphical format:

- Using inappropriate graphing technique such as using pie charts for non-mutually exclusive categories. Comparisons tend to be confusing when pie charts are used to plot non-exclusive categories. This mistake is less common among experienced researchers.
- Using pseudo 3-dimensional diagrams when the information is 2-dimensional. While the resulting graph may look pleasing to the eye, the meaning is completely distorted by this approach. Unfortunately this mistake is not confined to inexperienced researchers. As a matter of fact this is probably the single most common error made by all report writers.
- Not understanding what meanings are associated with different parts of a chart.
- Expecting the reader to guess the actual values involved (such as the value a bar of a given size represents).
- Presenting graphs that look impressive but are more difficult to read and understand-such as layered ('geology') charts.
- Presenting graphs when there are too many values that need to be accessed.

The list goes on. Unfortunately, while illiteracy is frowned upon and innumeracy is tolerated in our society, graphic illiteracy is not even recognized as a problem. So not much literature is widely available in this area.

A picture may be 'worth a thousand words' but a picture is meaningful only when it is appropriate and the communicator already knows what the thousand words are. An inappropriate picture is worth a thousand irrelevant words.

William Horton's book is an excellent introduction to this field of communicating visually. It explains when you should use mostly words, when you should use mostly graphics and when you should use both.

What is a graphic? Here is a definition.
Graphic = Message + Redundancy + Decoration + Noise

Good graphs concentrate primarily on message and redundancy while keeping noise to a minimum. Decorative purposes of a graph are relevant only in so far as the other functions are fulfilled. When we concentrate mostly on the decorative aspects, we do not notice the noise that it creates, obscuring and distorting the message of the graph.

Horton covers more than simple graphs. He covers the whole range of a graphic presentation. A particularly useful table in the book deals with when to use what. For example, what is the most appropriate chart for showing the correlation between two variables? If you want to show relative values, under what conditions would you use a pie chart? Under what condition would you use a bar or a column chart? When is a tree diagram appropriate? When do you use bullet points? When do you use numbering? What is the best way to present trend data? This book even covers the effective use of colour in charting. There is also a chapter on enriching graphics.

The amount of information that is packed into this book is indeed impressive. The book has wide margins so graphics can be presented in the margins with commentary. The text describes the principles of good and meaningful visual presentation. Another impressive feature of this book is an extensive bibliography of some 350 sources.

William Horton has provided us with an extensive, clear and detailed view of visual presentation. While the information presented is sophisticated, Horton's book is also a 'how-to-do' manual. If you have any occasion to present any information visually, if you are not clear as to when to use what type of chart (and why), if you always suspected that you could improve your visual presentation or if you believe that visual literacy is an important component of overall literacy, you can benefit from this book. Highly recommended.

Chuck Chakrapani

Dr Chuck Chakrapani is President of Standard Research Systems Inc. He is the author of several books and is the Editor-in-Chief of the Canadian Journal of Marketing Research.