Spring 2014, Volume 4, Issue 3

THE STATSWHISPERER

The StatsWhisperer Newsletter is published by Dr. William Bannon and the staff at StatsWhisperer.™ For other resources in learning statistics visit us on the web at: www.statswhisperer.com

The Effect Size: Advocating for a Measure of Magnitude

In prior materials we have mentioned that when examining the relationship between two variables, it is essential to note three relationship dimensions, which are significance, directionality, and magnitude (Bannon, 2013, p.50). Without considering each of these three dimensions, a data analyst cannot fully understand the relationship between two factors.

It is important to note that within the field of published research there tends to be an overemphasis on the first dimension, statistical significance. In fact within a quantitative study, identifying a relationship between variables with a probability level of .05 (p<.05) or less (i.e., achieving the most common criteria for statistical significance) commonly seems to be seen as the most important element within a study.

While statistical significance is important, it is in no way, shape, or form, more important than the third dimension, magnitude. However, as one peruses the literature, one cannot help but notice that this third dimension is often neglected or ignored entirely.

The current newsletter issue will take issue with this trend. In fact, after providing a brief description of what *magnitude* is and how it is measured, we will absolutely advocate for this relationship dimension. Why? Because if you cannot observe the magnitude of the effect one variable has upon another, you cannot interpret the variable relationship meaningfully. Within the current newsletter, we will examine magnitude through observing the effect size of the variable relationship, which is a common measure of magnitude.

One could suggest that observing issues related to *magnitude* represents a key in the evolution of quantitative research. In short, an equal emphasis on magnitude relative to statistical significance, will facilitate a richer and more complete interpretation of how factors relate to one another within a quantitative study.

INSIDE THIS ISSUE

The Three Dimensions of a Relationship

Here we will present a brief description of the three dimensions of a variable relationship:

1) Statistical Significance: In testing the relationship between variables, significance, or more specifically, statistical significance is expressed through the level of probability, a.k.a. the *p*–value. Probability reflects how likely the relationship presented is a result of chance. For example, let's suppose we surveyed 100 men and 100 women on their level of happiness using a scale from 1-5 (with 5 indicating higher levels of happiness). Let's also suppose that the males had a mean level of happiness of 2.5 and the females 3.5. Our probability level (p-value) would tell us how often we would get a difference like this by chance.

Although the level of probability may vary, the typical scientific standard is that probability must be less than 5 times in 100 (i.e., $p<.05$) to achieve statistical significance.

2) Directionality: Directionality indicates if a variable relationship is positive or negative. A **positive relationship** between two variables in statistics could mean, as the first variable increases in score, the second variable increases in score. For example, a positive relationship between the two variables age and happiness would suggest as age increases, happiness scores also increase. A **negative relationship** between variables suggests as the first variable increases in score, the second variable decreases in score. For example, a negative relationship between age and happiness suggests as age increases, happiness scores decrease.

Although there are several indicators, in statistical testing a negative relationship may be indicated by the presence of a minus (–) sign in front of the statistical coefficient. Also, a positive relationship may be indicated if there is not a minus sign (no –) in front of the statistical coefficient.

3) Magnitude: Once you indicate if a relationship is significant, as well as positive or negative, you might then wonder how strong the relationship is, which is a question of magnitude. A variable relationship may be statistically significant, but not very impactful or statistically significant and very impactful. A measure of magnitude must be used to determine which of these possibilities is indicated.

One of the most common measures of magnitude is the effect size. An effect size reflects the impact variables have upon one another and is often expressed as small, medium, and large. For the most part, each statistical test has a different value that reflects the effect size between variables. Although the raw numbers are different, all values have cutoff points reflecting a small, medium, and large effect size. For example:

Pearson's Correlation: The effect size for the Pearson's correlation is the either the r (small=.10, medium=.30, large=.50) or r-squared (small=.01, medium=.09, large=.25).

Chi-square: The effect size for the chi-square is the either the Cramer's V (greater than a 2X2 table) or Phi (2X2 table), both of which have an effect size of small=.10, medium=.30, large=.50.

It is critical to note that unless the magnitude, via a measure such as the effect size, is noted and understood, the true relationship between variables is somewhat hidden. In the next section we will illustrate how neglecting to report this factor leaves out a key part of the quantitative story.

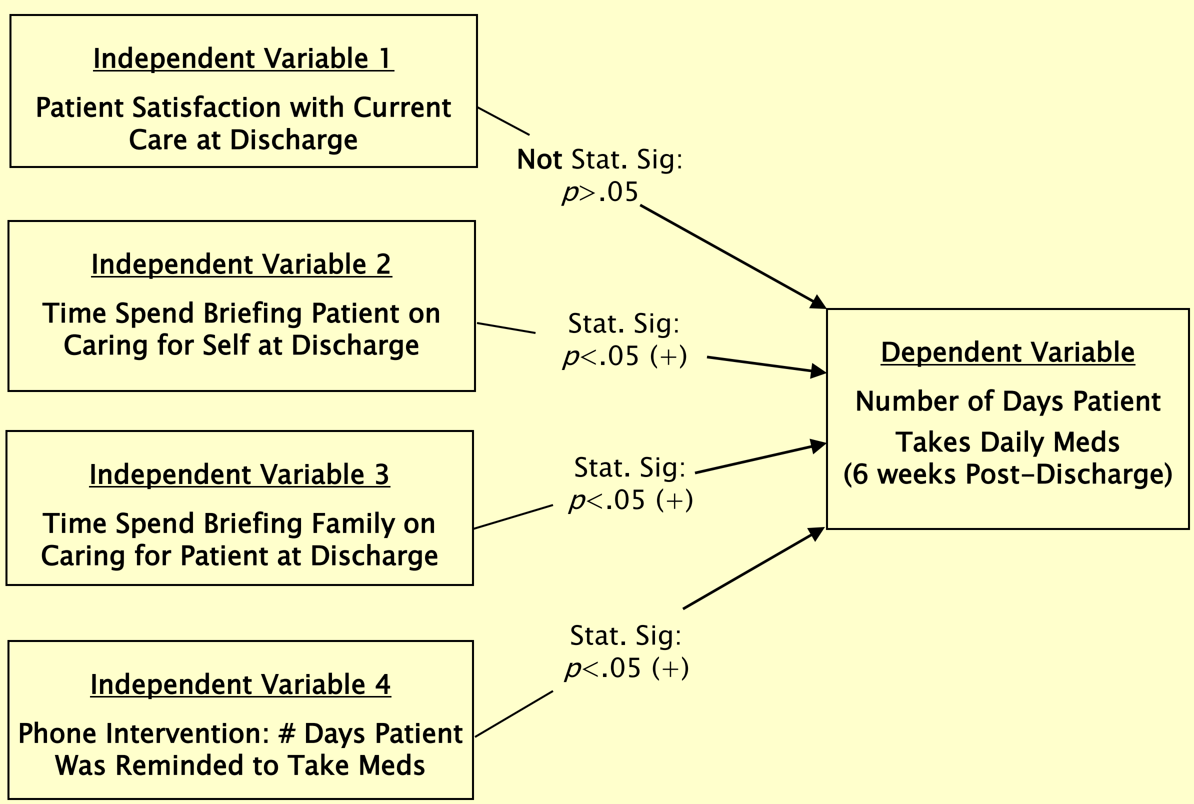
Reporting Statistical Significance Within a Study

Let's consider the most commonly reported relationship dimension within quantitative research, which is the level of statistical significance between variables. Specifically, consider **Diagram 1 The Level of Statistical Significance Between Study Variables**. In the center of the arrows connecting each independent variable (IV) with the dependent variable (DV) is the level of statistical significance (and directionality for significant IVs). Please note, all study variables (the IVs and the DV) are continuous. Subsequently, let's say that each arrow reflects the result of a Pearson's r correlation analysis, which would be appropriate to relate each IV to the DV.

Let's also assume that you are an administrator tasked with reporting which IV has the greatest impact on the DV, as that particular IV will become the focus of a new policy to be implemented.

When we observe the findings, we can see that there is not a statistically significant relationship between IV 1 and the DV ($p<.05$). However, we also see that there is a statistically significant relationship ($p<.05$), which we can see is positive by the plus sign (+), between each of the other IVs (IV 2, IV 3, IV 4) and the DV. This indicates that as each of these IVs increase, the DV increases at a statistically significant level.

Diagram 1. The Level of Statistical Significance Between Study Variables



Reporting Statistical Significance Within a Study (continued)

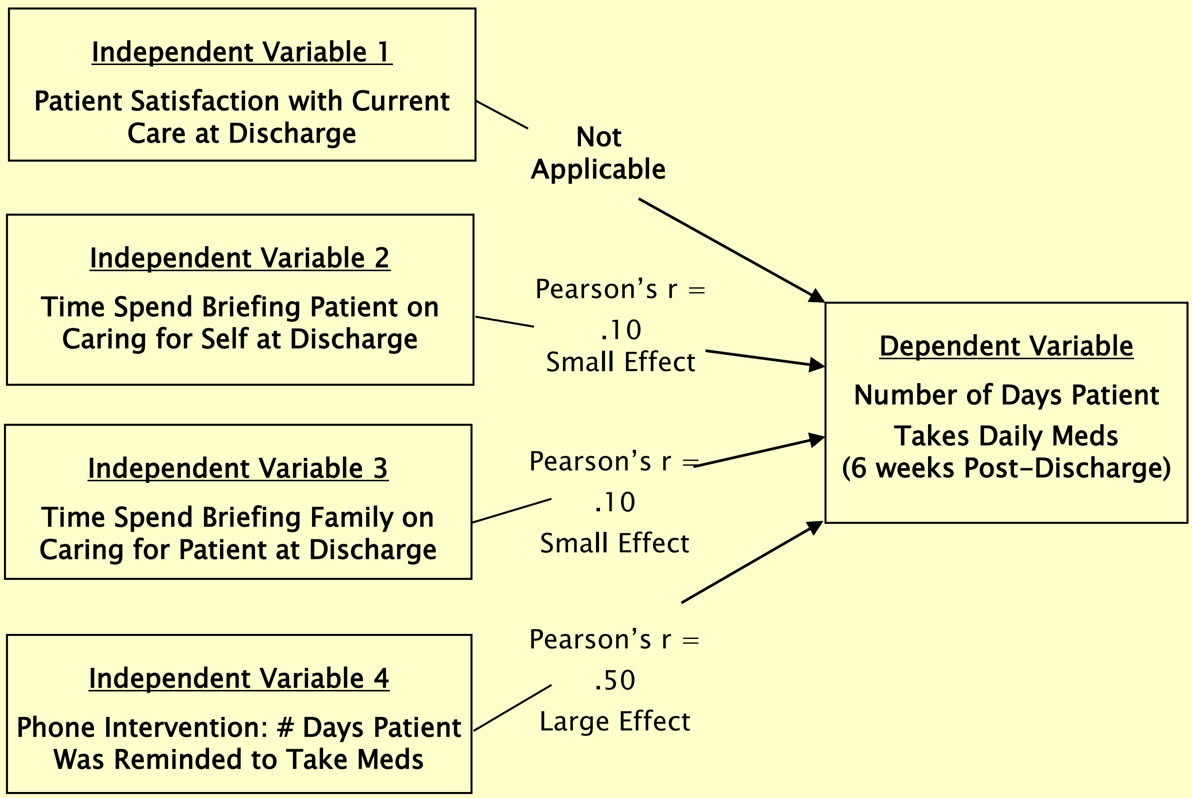
Doubtlessly, the presence or lack of presence of statistical significance is an important factor. To reiterate, statistical significance has now indicated that IV 2, IV3, and IV4 are significantly related to the DV, while IV 1 is not. This is the point where many quantitative studies leave off.

However, this is too little detail about the IV/DV relationships. To illustrate this, consider that this information is not sufficient to satisfy our task of identifying which IV has the greatest impact on the DV.

Now the question begs, just how can we make distinctions between each IV/DV relationship where we can identify the variable with the greatest impact on the DV among the three significant IV/DV relationships?

You may have guessed that these details will be illustrated using effect sizes to describe the magnitude of the impact each IV has upon the DV.

Diagram 2. The Effect Size Between Study Variables



Reporting Effect Size Within a Study

As stated earlier, the effect size describes the magnitude of the effect one variable has upon another. Please see **Diagram 2 The Effect Size Between Study Variables**. In the center of the arrows connecting each independent variable (IV) with the dependent variable (DV) is the Pearson's r correlation, which is the effect size reflecting the impact of the IV on the DV. Recall as stated earlier, a Pearson's r of .10 is considered a small effect size, a Pearson's r of .30 is considered a medium effect size, and a Pearson's r of .50 is considered a large effect size.

We can see that the arrow connecting IV 2 to the DV suggests that IV2 **Time Spend Briefing Patient on Caring for Self at Discharge** has a small effect (Pearson's r = .10, which is a small effect) upon the DV **Number of Days Patient Takes Daily Meds (6 weeks Post-Discharge)**.

We can see that the arrow connecting IV 3 to the DV suggests that IV3 **Time Spend Briefing Family on Caring for Patient at Discharge** also has a small effect (Pearson's r = .10, which is a small effect) upon the DV **Number of Days Patient Takes Daily Meds (6 weeks Post-Discharge)**.

However, we see that the arrow connecting IV 4 to the DV suggests that IV4 **Phone Intervention: # Days Patient Was Reminded to Take Meds** has a large effect (Pearson's r = .50, which is a large effect) upon the DV **Number of Days Patient Takes Daily Meds (6 weeks Post-Discharge)**. Thus, through observing the effect size (i.e., the Pearson's r) we can make a distinction between the three significant IV/DV relationships to identify which IV has the greatest impact upon the DV, which satisfies the demand of the administrative task.

The Two Diagrams and The Take Home Point

The two diagrams are presented to illustrate one primary take home point. First, the dimension of statistical significance presented in Diagram 1 is important and necessary. However, the dimension of magnitude presented via effect size estimates, is an entirely different dimension of the IV/DV relationship, but is also essential.

As stated earlier, there is a strong focus in the quantitative research literature on statistical significance, while measures of magnitude are often downplayed or not mentioned at all. Thus, there is a tendency to simply categorize IVs (and other predictors) as related to the DV at a statistically significant relationship *Yes* or *No*.

In short, when IVs are placed in the category of being related to the DV at a statistically significant level *Yes* or *No*, there is a suggestion that the IVs placed in the *Yes* category are pretty much the same. However, as we see in **Diagram 2**, IVs can all be related to the DV at a statistically significant level, but have very different relationships with the DV.

For example, IV 2 and IV 3 have small effects on the DV. As such, perhaps each does not even warrant consideration as a factor that might be incorporated within an intervention to impact the DV. However, IV 4 has a large effect on the DV, which might suggest that this is a

The Two Diagrams and The Take Home Point (continued)

a prime variable to target in order to enhance outcomes related to the DV.

However, if a data analyst only knew to consider the level of statistical significance between the IV/DV, he or she might have regarded each on the statistically significant IV/DV relationships as equal candidates to inform an intervention. This is a perfect illustration of how if the dimension of magnitude is neglected, a data analyst would have only a partial understanding of the IV/DV relationship.

Final Comments

The examples used in the current newsletter are rather simple. However, this is done purposefully to make our point clear. Namely, in every quantitative study it is critical to report measures of magnitude, such as the effect size, to more fully illustrate the relationship between variables.

As we stated, many studies treat IVs that are significantly related to the DV as equal. However, an IV can be significantly related to a DV, but be rather inconsequential if the effect size is tiny. Additionally, an IV can be significantly related to a DV and be incredibly consequential if the effect size is large. Therefore, it is critical to make further distinctions regarding IV/DV relationships beyond statistical significance.

The study of measures of magnitude is quite an extensive subject in itself. For example, as mentioned earlier, there is typically a different effect size measure for each type of statistical test. Within each effect size measure, the cutoff values reflecting small, medium, and large effects vary. However, it is not difficult to determine the proper effect size for each test (there are many online resources), as well as the cutoff values for small, medium, and large effects. **What is less common, is an awareness that the effect size exists and it is critical to include these estimates in each quantitative study. This awareness is the primary goal of this newsletter issue.**

In the beginning of this newsletter issue we used the term *advocacy* in reference to the inclusion of estimates of magnitude within quantitative research. This is critical, as while a fair number of researchers realize the importance of measures of magnitude, there do not seem to be a large group advocating for these estimates to be included in all quantitative studies.

However, as the number of researchers increase that are aware of and desire to observe estimates of magnitude such as effect sizes, in quantitative research, the more normalized the presentation of these values will become. In end this could represent a real step forward in the general approach to statistical research.

REFERENCES

Bannon, W. M. (2013). *The 7 Steps of Data Analysis: A Manual for Conducting a Quantitative Analysis*. New York: StatsWhisperer Press. Only available at:

http://www.statswhisperer.com/books/

Thank You for Your Interest In Our Newsletter!