AnalyticsIQ

Cognitive Psychology: The Missing Piece in Your Data Strategy

Understanding “how” and “why” your audience makes decisions.

IN THE DIGITAL UNIVERSE where business competition is fierce and advertising is relentless, marketers need every advantage they can get to reach the right customer with the right message. Many give themselves a better chance of doing exactly that by using big data and predictive analytics to anticipate the 'who', 'what', 'when', and 'where' of consumer actions.

Cognitive psychology is the scientific study of the mind, focusing on the way people process information and then understanding how this translates into behaviors.

Even though we live in the era of "big data", marketers often struggle to get their hands on information that also answers the question "Why?". Why and how do consumers make their purchase decisions? What motivates them to take action? As James Kobielus, IBM Data Science Evangelist states, "Most pop psychology that pervades marketing is shallow…and largely unsupported by solid academic research." With a perceived shortage of data based on true cognitive science, marketers are often missing a critical piece of their data puzzle, one that can help them better personalize consumer interactions, optimize campaigns and predict future needs.
To solve this problem, AnalyticsIQ has pioneered a sophisticated approach to curating data by pairing cognitive psychology with data science. In fact, we are so focused on understanding the 'why' that our team is comprised of industry veterans, data scientists and even cognitive psychologists. Our team is on a mission to get to the heart of the psychological motivators and predictors that fuel consumers' decisions.

"The science of psychology—why people are doing what they are doing—in traditional marketing research provides a great complement to what can be measured."

ERIC BRADLOW, PROFESSOR OF MARKETING, THE WHARTON SCHOOL

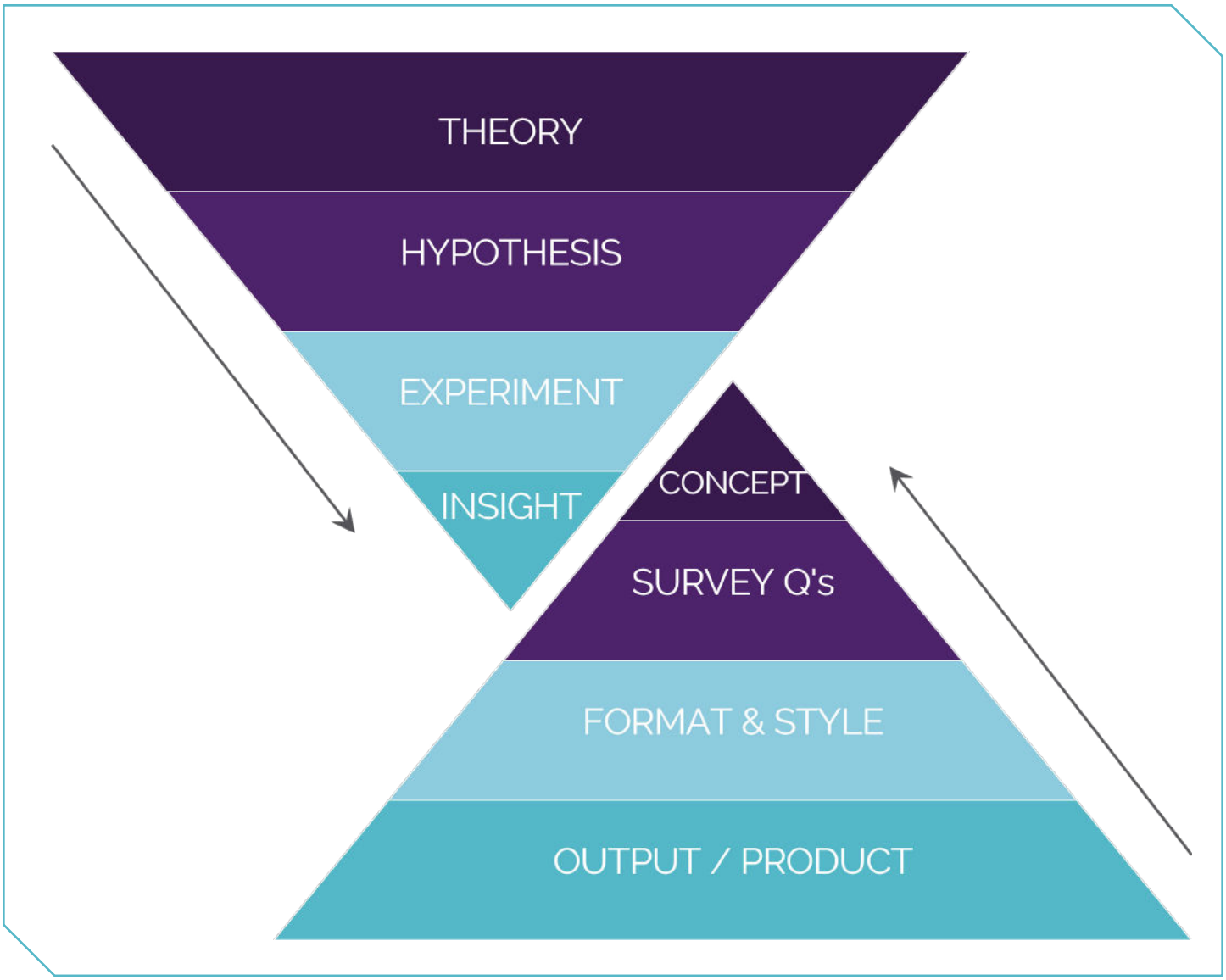
WHAT IS COGNITIVE PSYCHOLOGY?

The field of cognitive psychology is fundamentally the study of thoughts, thought processes, and mental mechanisms. It emerged in the late 1800s because of the foreboding need to take philosophical questions about human nature and the mind into a controlled setting where the phenomena of interest could be broken down into its component parts and rigorously studied. The field of cognitive psychology is responsible for the advancements in our understanding of topics like attention, memory, categorization, language, decision-making, and implicit biases, to name a few.

As with all other scientific disciplines, cognitive psychology utilizes the scientific method to gain a greater understanding of human thoughts and behaviors. Generally, the scientific method includes the following steps: identifying the research question, conducting background research, forming hypotheses, developing and executing experimentation (or assessment), conducting data analyses, drawing inferences and conclusions, and communicating the results. This 7-step process is the foundation for each project conducted at AnalyticsIQ.

HOW DOES COGNITIVE PSYCHOLOGY INFORM THE WAY WE BUILD DATA?

Our product development process combines the needs of the marketplace and our clients, expert knowledge of cognitive mechanisms, and the rigor of the scientific method, with predictive modeling analytics. The result is a hybrid model of top-down and bottom-up processes that take the following form:



1. **DEFINING OBJECTIVES:** The cognitive sciences team begins by starting all the way at the bottom in order to identify where we hope to end up. What do we want to build or know (e.g., data product, behavior, motivation, target population)?
2. **LEARNING AND EXPLORATION:** Then, we jump to the top and dive into the scientific literature to learn everything we can on the topic identified from Step 1. It's here that we identify any relevant theory to inform our research question and gain a broad understanding of any past research conducted on the topic.
3. **IDENTIFYING METRICS:** Next, we identify (a) key behaviors, (b) belief systems and (c) cognitive mechanisms we want to include in our survey (or experiment), along with the best tools to assess them. This could include developing scales, measures, and assessments of our own when one is not available to meet our needs.
4. **PROTOCOL DEVELOPMENT:** Then, we identify the best methodology (e.g., survey, experiment) to achieve our goal, prepare the question-by-question protocol, and program the protocol in preparation for data collection.
5. **DATA COLLECTION:** Next, we gather the data from the desired population. In most cases, this is a representative sample of the US population (based on age, gender, race). Other times, our research samples are niche to a particular client and their consumer base or to a specific population segment (e.g., demographic). We typically work with outside vendors to facilitate this stage of the research process.
6. **DATA CLEANING AND EVAULATION:** Finally, we bring the data in-house where it undergoes a meticulous process of data cleaning and evaluation to ensure that only the most pristine data are included in the analyses that follow.
7. **DATA ANALYSES:** Once the Cognitive Sciences team has done all of this, then the raw data moves down two complementary pathways.
    a. First, it is sent to AnalyticsIQ's team of data analysts who then develop the predictive data models. Once our Data Analytics team completes a predictive model, it then goes through an established validation process before being ready for inclusion in our PeopleCore or BusinessCore databases.
    b. Second, it is evaluated and analyzed by the Cognitive Sciences team according to the project goal and overarching research question (e.g., exploratory analyses, evaluating a priori expectations, hypothesis testing, qualitative coding). The results of these analyses are then written up into a comprehensive research report or compiled into a quick and digestible flipbook for industry consumption and insight.

WHAT MAKES OUR DATA DIFFERENT?

Reliability and validity are key characteristics of the empirical research process, and of our data at AnalyticsIQ, because they ensure that the necessary checks and balances are built into our research and data development processes. Construct validity, content validity, and criterion validity are top of mind for our cognitive scientists when building measurement and assessment tools. Internal and external validity are critical considerations when evaluating predictive (cause-and-effect) relationships. And test-retest reliability, along with representativeness, are foundations to all our research projects no matter the structure (survey, experimental) or purpose (data product development, theoretical insights). Let's take a look at each one of these criteria individually to see how they impact the quality of the data we report.

* ***Construct validity.*** Does the assessment measure the construct that it is intended to measure?

An Example:

In 2019, we wanted to develop a predictive model of people who had the highest likelihood of experiencing good (vs. bad) quality sleep. In order to have strong *construct validity* we first needed to identify and operationalize the components of what constitutes "good sleep". The research literature indicated that having healthy sleep patterns was more complex than the number of hours a person reported sleeping or even a subjective rating of their sleep quality. Rather, we discovered that the best assessment of "good sleep" was a combination of (1) subjective sleep quality, (2) # of hours of sleep, and (3) the presence (or absence) of sleep disturbances. These three components were critical elements that needed to be included in our sleep assessment to meet the construct validity requirement.

- ***Content Validity.*** Do the items within the assessment adequately cover all components of the desired construct?

An Example:
Next, we had to ensure that all three sleep construct components (subjective quality, average hours, sleep disturbances) had *content validity*. In doing so, we discovered that properly assessing subjective sleep quality required us to ask participants about feelings of tiredness upon waking up and feelings of tiredness throughout the day (rather than simply asking people to rate how "good" or "bad" their sleep was on a 1-7 likert scale). We also discovered that the presence of sleep disturbances was a complex combination of various nighttime experiences, including:

- Trouble falling asleep,
- Trouble staying asleep,
- Night sweats,
- Nightmares,
- Use of an ingestible sleep aid (e.g., pharmaceutical, herbal supplement, alcohol, cannabis).
- Among others…

Once we had identified the best *content* to include in our sleep assessment, we were able to run statistical analyses to validate our assessment tool (e.g., factor analysis). This then informed the creation of a quantifiable score of "good sleep" which then became the foundation for our HW_Sleep model.

Note: If our analyses had revealed that multiple factors were generated from the items in our assessment or that certain items were not statistically interrelated with the other items, then this could signal that our measurement lacked content validity. In this case, we would need to adjust our assessment tool guided by these statistical indications.

- ***Criterion Validity.*** How does the assessment perform relative to other separate (but related) measures and their known relationship?

An Example:
Sleep has a documented relationship with depression, anxiety, and stress, such that an increase in sleep difficulty is significantly correlated with an increase in symptoms of depression, anxiety, and stress. To ensure *criterion validity*, we ran a simple bi-variate correlation between our sleep and stress assessment scores and found a significant positive correlation between the two measures as anticipated. This is one indicator that we have met criterion validity.

- ***Internal Validity.*** Is there a cause-and-effect relationship present among the key variables, and if so, can that relationship be attributed to anything other than the key variables? In the case of an experiment with multiple levels of an independent variable (e.g., treatment vs. no treatment), are the effects observed in the experiment due to the various levels of the independent variable? The best measures of evaluating *internal validity* are statistical significance (p-value) and effect size (Eta^2, R^2).

The p-value (or 'probability value') ensures a critical level of *confidence* that the effect you are observing in the data is due exclusively to the variables of interest. A p-value of <.01 means you are 99% confident that the observed effect is due to the variables in the statistical model and not extraneous variable(s) or statistical error. P-values outside that threshold would not meet the minimum requirement to be considered 'statistically significant' (e.g., $p<.03$).

The effect size, on the other hand, measures the *magnitude* of the observed effect. Two common metrics of effect size are Eta^2 and R^2. Eta^2 is used most often when running Analyses of Variance (ANOVAs), whereas R^2 is most common in regression-based analyses. Both provide a measure of the proportion of variance in the outcome (i.e., dependent variable, DV) associated with each independent variable (IV). Effect size metrics serve to complement the p-value, such that the internal validity of your project is based on a combination of confidence (p-value) and magnitude (effect size). Internal validity has been met when both metrics are above minimum thresholds.

An Example:

To ensure *internal validity*, we ran a simple linear regression between sleep and stress assessment scores and found a significant predictive relationship between the two measures (as was found with the bi-variate correlation). However, in this case, we used a cause-and-effect analysis and had both a confidence level and an effect size to weigh against one another. Our statistical significance was $p<.001$ (or 99.99% confidence) and our effect size (R^2) was .25 (or 25% variance explained). These metrics combined offer confirmation that our assessments (and the relationship between the constructs) have internal validity.

- ***External Validity.*** Can the measurement tool or experimental effect generalize to other contexts or settings? The best way to ensure *external validity* is to test the initial observed effect across populations or contexts that are different from the original assessment or study.

An Example:
One way to examine the *external validity* of our Sleep Assessment was to evaluate the observed relationship between stress and sleep across folks within various age brackets (e.g., Gen Z, Millennials, Boomers), gender identities (e.g., Male, Female, Trans), or even geographic regions (e.g., Northeast, Southeast, Mid-West, West; Domestic vs International). Doing so and finding the original relationship across each population segment or context would suggest that the measure and / or effect was generalizable and thus had external validity.

When we examined the stress-sleep relationship across age brackets and gender groups we indeed found the same predictive relationship that we found previously (i.e., higher stress predicts greater sleep disturbances; all segments showed statistical significance, $p<.01$). However, there was some unanticipated variability in effect sizes across each demographic segment (R^2s ranged from 21% - 28%). Taken together, these data suggest that not only are our assessments of sleep and stress externally valid but that the predictive relationship between them is also externally valid. Moreover, there is some meaningful variability in how strong this stress-sleep relationship is across population segments, which provides inroads for follow-up research exploration.

- ***Reliability.*** Is the measurement, assessment, or experimental effect consistent over time? This is also known as test-retest reliability. The best way to test for test-retest reliability is to collect longitudinal data and evaluate your assessment tools or experimental effects across multiple time points.

An Example:
To ensure data and measurement *reliability*, data for our Sleep Assessment was collected at a second time point with a new sample of participants. Once the data was collected and cleaned, two analyses were conducted: a confirmation factor analysis (CFA) and a linear regression. The CFA was conducted to rebuild the composite score from time point 1 in order to ensure that the items included in the original score were statistically inter-related in the same way at the time point 2, and the linear regression was conducted to ensure that the causal relationship between stress and sleep was also present at time point 2. Model fit statistics from the CFA confirmed that our assessment was valid, and the linear regression once again confirmed the presence of a significant positive relationship. Taken together, these metrics suggest strong test-retest reliability.

- ***Representation.*** Does your sample represent the population of interest? In almost all cases, our intent is to sample a subsection of US adults that portray national averages on age, gender, and race (based on US census data). Some segments (e.g., Latinx, LGBTQ+) are naturally tougher to find in the online sampling pool, which means that we are in constant communication with our sample vendors to improve our recruitment methods so that we continue to have a diverse presence in every one of our research datasets.

LET'S TALK ABOUT ERROR

Any scientist, analyst, or mathematician will tell you that some form of error will always be present because measurements and assessments are only ever an approximation of the true value of a behavior, motivation, or intention. In fact, error (*e*) is commonly built into mathematical and statistical formulas for this exact reason. Thankfully, there are predictable places to look for error and reliable approaches to minimizing these gaps in reliability or validity. Some key sources of error include participant inaccuracy, invalid measurement tools, and human error when processing data.

Participant Inaccuracy

When we discover inaccuracy in respondent data, it's rarely because research participants are intentionally being untruthful when answering questions. In fact, where there are gaps in accuracy, most often it is because of one or more of the following:

1. Participants aren't motivated to be honest and accurate.
- To increase motivation, all participants are incentivized with financial compensation for their participation. Longer research protocols (or protocols that require qualitative responses) offer higher compensation.

2. Participants are moving through the survey as quickly as possible to get to the end and collect their reward. We call these folks "speeders" and "straight-liners."
- To identify speeders, we typically look at the total amount of time it took them to complete the research protocol (i.e., runtime). If the total runtime took less than a minimum threshold (commonly half the average runtime), then we have reason to "flag" that participant and examine their data more closely.
- Straight-liners are aptly named because they speed through the survey by entering the same response (e.g., "1") for every question. This, thankfully, makes them easy to identify. To catch straight-liners, we spot-check sets of questions by sorting responses low to high and then look for participants who gave the same answer to a consecutive set of questions. Straight-liners are flagged, and their individual data examined more closely before deciding whether to exclude or include their data in follow-up data analyses.

3. Participants are not paying attention or reading the questions fully.
- To identify inattentiveness, we include "attention checks" in the research protocol (e.g., "If you're paying attention, please type the letter 'd' and then hit enter," "Please identify the colors of the American flag from the list below"). Attention checks always have a right and wrong answer, and we usually place 2-3 of them throughout a protocol (depending on overall protocol length; shorter surveys require less attention checks).

4. Participants might not understand the question being asked.
- To ensure our questions are simple and clear, we write all our protocols at an eighth grade reading level. This is a common standard across institutional research bodies and ethics review boards.

5. Some level of self-knowledge is required to accurately report your own behavior, belief, or motivation, and there is great variability in one's ability to introspect in these ways.
- To identify inaccuracy based on lack of self-knowledge or self-assessment, we will commonly ask the same question multiple times but use different verbiage (and / or response choices) so that we can validate those questions against each other.

Invalid Measurement Tools

The section above on reliability and validity largely outlines how we attempt to minimize error in the assessment and measurement tools we build ourselves. Please refer back to that section again as needed. But what about measurement and assessment tools that other scientists or researchers have built? Thankfully the same standards around reliability and validity exist for all scientists, and publishing one's work in high-impact journals requires a rigorous review process by multiple highly qualified and specialized scientists to ensure that all critical standards have been met. When a measurement tool is identified for use in one of our studies, our team of cognitive scientists ensures quality control by conducting analyses (e.g., Cronbach's alpha, CFA) to ensure both reliability and validity.

Human Error When Processing Data

We are just humans after all, and all humans are fallible. Because we are humble enough to know this, we put in place our own departmental practices intended to catch the most common human data errors. These practices include inter-rater reliability, spot-checks, template creation, and automation (where possible). Inter-rater reliability occurs when two or more independent "raters" (research scientists) agree on a subjective coding scheme for a set of data. There is typically a minimum threshold of agreement that must be met to achieve this. Spot-checking is simply the process of randomly re-running a subset of a colleagues' analyses in an attempt to catch errors or inaccuracies. Developing analysis automations and reporting and formula templates can greatly reduce your chances of human error.

OUR COMMITMENTS TO YOU

Transparency and ethics are the backbone of the cognitive science research process. We are committed to ensuring that every person who participates in a research project with us does so only after fully understanding what we are studying and why. As such, all participants must review and agree to our Informed Consent. The informed consent document includes details about the purpose of the project, how their data will be stored once their participation is complete, and confirmation that their responses will be de-identified from any personally identifying information. Participants must affirm that they understand and agree to the terms presented in the informed consent prior to joining our research studies. Moreover, all participants are compensated for their time, and there is never any penalty for choosing to abstain from answering any question in any of our research protocols. Furthermore, transparency in methodology and data analyses is a critical foundation for building trust in our data and in our partnerships. Although some details about protocol items or question verbiage may not be shared directly (for proprietary reasons), our processes, methodologies, data analyses, and statistical results will always be communicated fully and accurately. The Cognitive Sciences team takes these commitments seriously.

SUMMARY

- Cognitive Psychology is the study of thoughts, thought processes, and mental mechanisms. It is responsible for advancements in our understanding of topics like attention, memory, decision-making, to name a few.
- As a scientific discipline, cognitive psychology utilizes the scientific method to gain a greater understanding of human thoughts and behaviors.
- The research and product development process at AnalyticsIQ uses a hybrid model of top-down and bottom-up processes.
- Our cognitive scientists take meticulous effort to ensure data reliability and validity, year over year.
- Error, or data inaccuracy, are a reality that all scientists and analysts face, and our team of cognitive scientists make every effort to minimize and mitigate all barriers to accuracy.
- We are committed to conducting ethical scientific research and transparently reporting our methodologies and data insights.

GET IN TOUCH. WE SPEAK GEEK.

For more information on these Social Relationship research insights or to inquire about conducting a custom research project of your own, reach out to **sales@analyticsiq.com**.

About the Author

Dr. Sarah Cavrak

Sarah Cavrak, PhD is a Psychologist, and the Senior Director of the Cognitive Sciences Department at AnalyticsIQ. She has spent 20 years studying psychological underpinnings of human behavior, and is primarily interested in understanding the intersection between motivational dynamics and decision outcomes. **Follow her on LinkedIn**.

AnalyticsIQ