Chapter 8
The Hypothetico-Deductive Method

8.1 Introduction

This is where the researcher stands right now (Fig. 8.1):

Once the researcher walks down the hypothetico-deductive path, constructs and bubbles are called variables.

It is now time for him to formulate some hypotheses regarding his model. He should always keep in mind, however, that in real life he sometimes proceeds by making assumptions and formulate hypotheses, but he also relies on gut feelings and induction. So a step in the hypothetico-deductive world is NOT the end of the world or of his research: it is part of his journey to becoming an expert in data percolation.

8.2 Types of Research

The researcher should not work on hypotheses before he is reasonably satisfied with his model. Once this has happened, he wants to confirm what kind of research he wants to do. He may have originally planned to restrict himself to descriptive research (finding a definition for his constructs—for example, a definition for a construct such as perceived predation using (*S*) and (*F*) bubbles), but he may also want to find out more.

There are essentially four types of research. Unfortunately, some researchers tend to confuse them and claim, for example, that they did causal research when in fact all they did was pure speculation.

O. Mesly, *Creating Models in Psychological Research,*
SpringerBriefs in Psychology, DOI 10.1007/978-3-319-15753-5_8

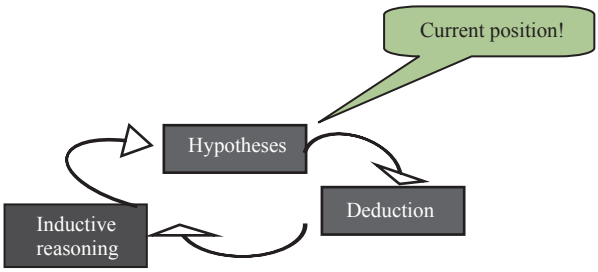


Fig. 8.1 A road map to evolution—Part II

Here are the four types—choose one:

- Descriptive (which can be comparative)—one uses (*S*) and (*F*) arrows
- Relational (measuring the influence of one variable on one or others)—one uses the (*I*) arrow or else chained arrows (*H*)
- Predictive (longitudinal)—one uses the (*T*) arrow
- Causal—one uses the (*C*) arrow

It is true that during a hypothetico-deductive investigation, the researcher should try to be as objective and as unemotionally "connected" to the respondents as possible. In an ideal world, one could test the variables in highly controlled or fully controlled circumstances. The researcher would have two clone groups, one where one variable would be tested, and the other where it would not, and all the other variables would be held constant, just as is attempted in research labs around the world. This is often impossible to do in the area of social sciences and psychology, in particular, which is one of the reasons the researcher wants to learn as much as possible about the different significant variables using literature review, expert opinion, qualitative studies, and possibly simulation prior to conducting hypothetico-deductive research. It is also one of the reasons he looks for contrasting cases. In reality, there are few controlled variables in psychology, hence the requirement to resort to data percolation.

An example of a descriptive study is that of Levenson (2004) where two types of criminals were compared on a series of preset parameters: those initially selected for release and those actually released in the end. Parameters included previous treatment failure use of weapon or infliction of injury, documented or admitted history of variety of sex offenses, history of murder, or attempted murder, and so forth. This descriptive research allows the researcher to argue that "The results provide preliminary but encouraging data suggesting that the highest risk sex offenders are being appropriately selected for commitment" (p. 646).

In the four different research types discussed above, the "I" arrow points to the fact that one variable influences another. The antecedent variable is an explanatory variable and the consequent variable is an explained variable. An example of influence is the role of gestural misinformation in skewing eyewitnesses' testimonials of crime scenes. It has been found that eyewitnesses are not only influenced by verbal cues but also by nonverbal cues (Gurney et al. 2013).

The influence can be direct or indirect (moderating or mediating variables) positive ($I+$), negative ($I-$) or else positive–negative ($I\pm$ for moderating variables

only). The influence involves time, but time can be very limited; for all intents and purposes, it is possible to have bubbles that are vertically positioned in a model with "I" arrows in between them: one recognizes that the influence of one variable on the other is nearly simultaneous. More on this will be seen below.

The "T" symbol indicates that there is a clear time factor between the bubble on the left and the bubble on the right (the flow is always left to right with "T"); however, it does not mean that the left bubble (independent variable) is the cause of the one on the right (dependent variable). As mentioned before, some authors make the mistake of confounding influence and cause-and-effect relationships. This kind of "T" study is also called longitudinal or predictive, because often it can help the researcher predict what will happen if the current trend persists while other variables are held constant. An example where longitudinal study is required is in the case of the assumption that early childhood attachment distress leads to an adult's tendency to depression; only by looking at a large number of people from the time they are children (and suffer attachment distress) to the time that they are adults (and suffer from a tendency to depression or from depression) can a valid answer be provided (Morley and Moran 2011). A study by Sutin et al. (2011) with $N=4790$; age range 14–94 is not a longitudinal study per se, because it is not the same group of respondents that were analyzed throughout their lifespan from 14 to 94 years of age. Rather, 4790 people ranging in age from 14 to 94 years are assumed to be equal so that the 94-year-old respondent is assumed to be a good representation of what the 14-year-old respondent would be at 94. Based on this assumption, the authors are able to posit that personality traits ("prospectively") predict verbal fluency.

To conduct a longitudinal study, the researcher must measure a phenomenon at one point in time, hold his breath in the hope that no exogenous variables (so-called "externalities") come and affect the participants, and measure the phenomenon once more later. He can trick the time factor by measuring two similar groups of participants that do not influence each other, but for which group 1 goes through the phenomenon as it would exist at the point of time 1 and group 2 goes through the same phenomenon as it would exist at the point of time 2. This is nearly impossible to do.

The causal "C" arrow is the most difficult one to study. It is represented as follows (Fig. 8.2):

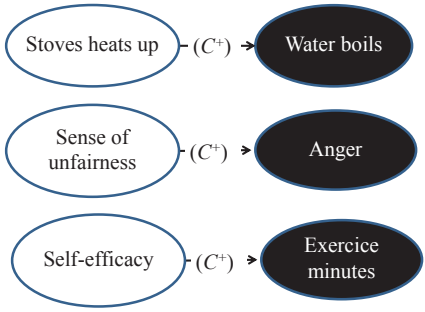


Fig. 8.2 The causal arrow

Table 8.1 The research's level of difficulty

	Type of research	Level of difficulty
Structural (*S*) and functional (*F*)	Descriptive	Easy
Influence (*I*)—positive or negative	Relational	Challenging
Longitudinal (*T*)	Predictive	Difficult
Causal (*C*)—positive or negative	Causal	In psychology, nearly impossible

There is no choice for the bubble on the right: it will necessarily occur given the bubble on the left (unlike a longitudinal study), 100 % of the time. Given a specific level of atmospheric pressure, water that is heated starts evaporating. Heat causes the water to evaporate: it will always evaporate given the appropriate heating level. The effect can be positive or negative; for example, some researchers state that smoking causes lung cancer, which is a negative outcome (yet others fight this conclusion vehemently). It is an error to have double-headed arrows (or chained arrows) with causal relationships.

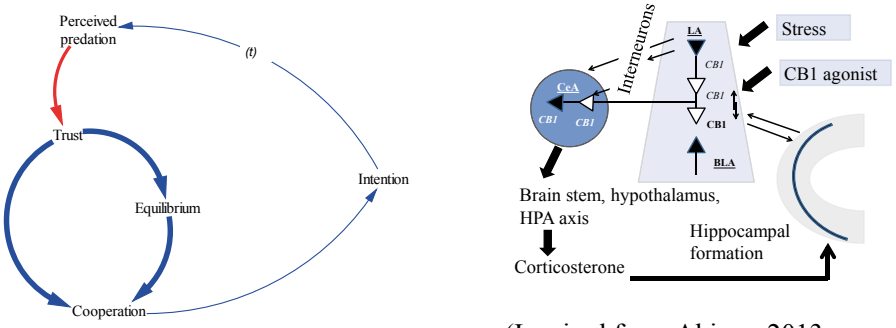
One way of convincing oneself that anger is caused by a sense of unfairness is to look at other kinds of populations, where this always occurs. For example, Seymour et al. (2007 p. 306) observe that "… chimpanzees attack allies that do not support them in third party conflicts, and queen naked mole rats will attack workers that they judge to be lazy". This may not be a proof but a strong indication of the assumption that chimpanzees realize that there is unfairness and that they react upon this realization. In their study, Basen-Engquist et al. (2013, p. 1137) resort to the concept of causal relationships as follows: "The consistency of the relationship between self-efficacy and exercise minutes over short (same day) and longer (Tj–Tj-1) time periods provides support for a causal relationship." Thus, causal links have found a place in psychological studies. Golden et al. (1987, p. 5) note: "Similarly, the cognitive-behavioral hypnotherapist assumes a direct causal link between cognitions or self-suggestions and emotional and behavioral consequences." Hence, here a causal link is assumed to take place.

The difficulty in doing research can be classified as follows (Table 8.1):

It is possible to add *retroactive arrows* (*t*) to one's model (loops) as follows (Fig. 8.3):

In Fig. 8.3, perceived predation diminishes the ability of the individual to trust others. A sense of unfairness (equilibrium) jointly with a lack of trust leads to lower social integration (cooperation) which may cause an intention to become violent, which then has a retroactive loop to the perceived predation construct.

Akirav's model (2013, p. 2560) is provided as an example. This is a typical model with retroaction because the pituitary gland does not feed back to the hypothalamus: hormones emanating from the hypothalamus or the pituitary gland must go through the blood stream before going back to the hypothalamus. In this example, stress affects the lateral amygdala (LA) and the basolateral amygdala (BLA), which then sends information to the central amygdala (CeA) and then the hypothalamus (HPA axis). The corticosterone travels through the blood stream to go back to the brain—the hippocampus and the amygdala. Borrowing from the constructs on the left of



(Inspired from Akirav, 2013, p. 2560)

Fig. 8.3 Retroaction (example) (This figure was obtained using Vensim as opposed to PowerPoint. Hence, different software will produce different ways of expressing the same model.) (Inspired from Akirav 2013, p. 2560)

Fig. 8.3 (which is a generalization of this process), this can be roughly expressed as perceived predation (stress) affecting trust (fear/amygdala), then affecting equilibrium (cognitive functions—in this case, memory/the hippocampal formation), with a retroactive loop eventually going back to trust (fear/amygdala).

The researcher will use the small "t". The retroactive arrows cannot go to a structural or functional bubble: these are timeless. They only emanate from and go to consequent and antecedent bubbles (variables).

8.3 Mediator and Moderator

Mediators ($I+$ or $I-$) and moderators ($I\pm$) are two forms of influence (I) arrows. Unlike the normal "I" arrow, they have an indirect influence on the variables they are in contact with. As put forth by Tofighi et al. (2013, p. 290) "Mediation analysis is a statistical approach used to understand how an independent variable produces an indirect effect on an outcome through an intervening variable (mediator)."

Figure 8.4 The mediating variable one can get to the right-hand bubble (B) by taking a direct road from the left-hand bubble (A). Alternatively, one could pass by the top bubble (Z) when one departs the bubble on the left (A) in order to get to the bubble on the right (B). An example of such dynamic is part of Moskowitz et al. model (2013, p. 1022) whereby recent stress acts as a mediating variable (Z) between emotional stress and suicide attempts. According to this model, emotional stress may lead directly to suicide attempts, but the presence of recent stressful events provides an alternative route that seems to encourage suicidal attempts. Chorpita and Barlow (1998, p. 9) propose that vulnerability acts as a mediating variable in their model on anxiety.

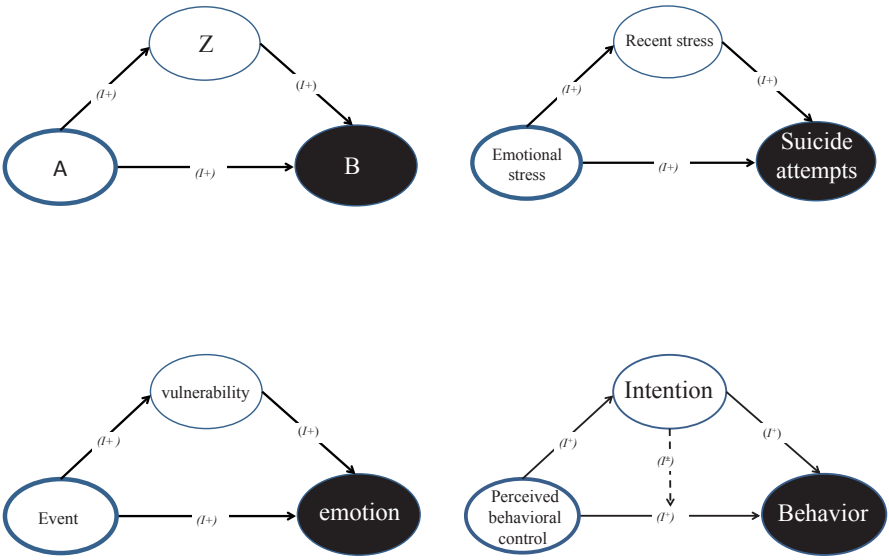


Fig. 8.4 Displays the mediating variable

As put in the example at the bottom right of Fig. 8.4, intention seems to be both a moderating and a mediating variable. Intentions are described by the authors as follows: "motivational factors that influence a behavior" and stronger intentions are associated with greater likelihood of performance or avoidance of a specific behavior in accordance with intentions (Ajzen 1991, p. 181). Thus, "individual's intentions to perform or abstain from a behavior are theorized to directly predict later behavior." (Rhodes and Clinkinbeard 2013, p. 26). However, a variable cannot act as both a moderator (a factor) and a mediator towards the same constructs, although a variable could act as a moderator for a set of constructs and as a mediator for a different set of constructs. In this case, clearly, intention is a variable of influence but not a factor in the sense of a moderating variable.

Baron and Kenny (1986) have developed an excellent technique for determining whether a variable is mediating or not. It is widely used. No qualitative study could really help the researcher decide if a variable is a mediating one, only a quantitative study can.

In the author's emerging consolidated model of predation (CMP), a vast array of participants and groups of participants were tested. It became clear that the construct "equilibrium" (win–win) was a mediating variable between trust and cooperation. This is how to interpret a mediating variable (example): a certain amount of trust could help develop cooperative efforts at the beginning of a relationship between a patient and his psychotherapist. However, if by some good fortune, each one senses that the encounter is a win–win situation, this may help or speed up the transition from the feeling of trust to cooperation. Equilibrium (win–win) is a mediating variable. One can live without it, but if it is there, that is good. One can go from point A

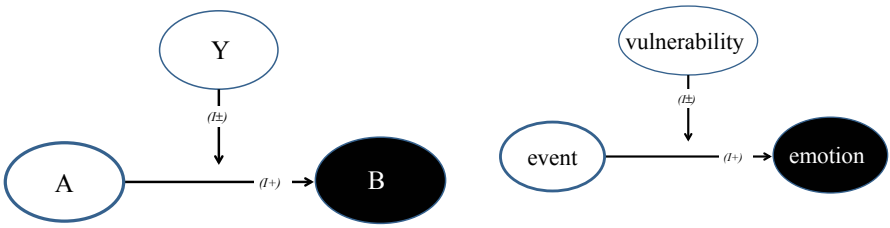


Fig. 8.5 A moderator

to point B without filling up one's gas tank, or one can go through point Z and get the opportunity to fill up one's gas tank at Z's gas station. In either scenario, one will end up at point B, but by going through point Z, one arrives at point B with a tank full of gas, which takes away the stress of having to find a gas station near point B. Point Z is a mediating variable.

There is also the moderating influence ($I\pm$). Figure 8.5 tells us a little bit about it:

It may be that the psychotherapist's personality (A) has a strong influence on the patient's intention to remain in therapy (B), but then, because the office where the therapy sessions take place is filthy, noisy, or has poor air conditioning (Y), it changes the patient's mood, despite the therapist's best efforts. The patient even decides to step out where there is fresh air. Generally speaking, a moderating variable is a factor that is external to the situation or the dynamics between the individuals. The best way to establish the existence of a moderator is to see how, for example, the participants behave with and without it (e.g., when the therapist re-establishes the flow of air conditioning (or cleans up his office), are the patients more eager to attend their therapy session?). In the example on the right, Chorpita and Barlow (1998, p. 9) propose an alternative to their initial model, with vulnerability acting as a moderator instead of a mediator. Moderation is statistically proven by a triangular distribution.[1]

Moderators are generally factors. In the field of criminal psychology, for example, the following external factors are thought to influence some individuals in becoming delinquent (while some others will seek to not fall into violence and compensate by excelling in society): culture (Fabrega 2004) and peer association (Katz and Marquette 1996). Applebaum et al. (1998), as another example, posit that conflict between individuals are shaped by both internal and external environmental factors.

One can also use statistical packages such as Partial Least Squares (regressions) (PLS). This is not always possible, of course, so to determine that the variable is an external factor is a good clue that it is a moderator. Moderating variables can be detected because they always lead to two opposite groups of reactions: in the case of the therapist's office, some patients will hurry to leave it because they want to escape the heat in it, while others will feel comfortable and secure, and even delay their leaving the office when the session is over. Hence, statistically, a moderating

[1] See Mesly and Lévy Mangin 2013; Mesly and Maziade 2013.

Table 8.2 All the arrows

Type of bond	Type of arrows
Structural (*S*) and functional (*F*)	Descriptive
	Binary (*Sb, Fb*)
	Continuous (*Sc, Fc*)
Influence (*I*)—positive or negative	Relational
Direct (*I* + or *I* −), chained (*H*)	
Indirect … mediator (*I* + or *I* −)	
Indirect … moderator (*I* ±)	
Longitudinal (*T*)	Predictive (*T*) or *(t)*
Causal (*C*)—positive or negative	(*C* + or *C* −)

variable has a triangular distribution (hence the use of the *I* ± sign). Many scientific papers do not recognize moderating variables and arrive at contradictory results; if the authors realized their results were contradictory because the variable was a moderating one, the debate would be closed. Note that in neuroscience, the concepts of moderating and mediating brain areas are used, but in a somewhat different sense.

Chained variables are two variables that influence each other concurrently, like two knights on critical squares of a chess board. This is found, for example, in the case of obesity and comorbid symptoms. Obesity leads to comorbid symptoms and comorbid symptoms lead to obesity. Similarly, it is generally recognized that PTSD is accompanied by comorbid manifestations, such as depression, drug abuse, social phobia, and so forth.

Let us summarize all the different types of arrows one finds under the data percolation methodology with Table 8.2.

That is all the researcher needs to create his model, yet many scientific models are, sadly, erroneous in their structure and explanatory power.

8.4 Hypotheses

In order to arrive at the tentative determination of *S*/*F* bonds, or *I, T*, or *C* arrows, the researcher must have been diligent in his modeling effort. Under the data percolation methodology, he must identify the type of hypothesis he is generating: $H^{(S)}$, $H^{(F)}$, $H^{(I)}$, $H^{(T)}$, $H^{(t)}$, or $H^{(C)}$. Also, a hypothesis can be tested according to two alternatives: H_0 (the so-called "null hypothesis") and H_a (its contrary). If the researcher has more than one hypothesis, each one should nevertheless be examined in consideration of the two options (each entailing some errors—type I and type II errors—consult with books on the subject). He must anchor his hypotheses in one of the four arrow modes (*S*/*F, I, T*, or *C*).

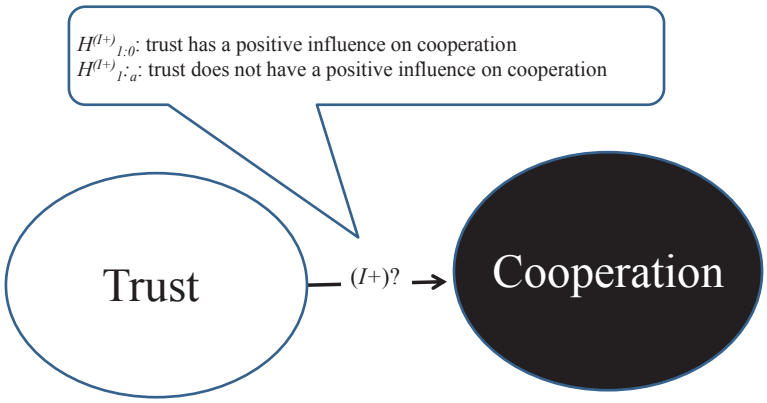


Fig. 8.6 Initial hypothesis. $H^{(I+)}{}_{1:0}$: trust has a positive influence on cooperation. $H^{(I+)}{}_{1:a}$: trust does not have a positive influence on cooperation

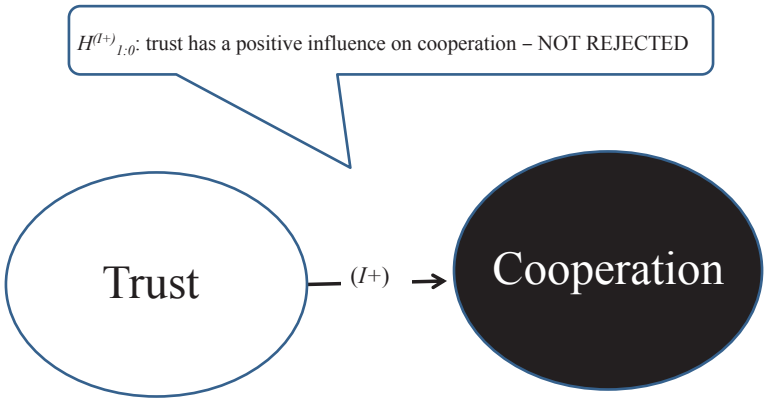


Fig. 8.7 Initial hypothesis with results. $H^{(I+)}{}_{1:0}$: trust has a positive influence on cooperation. $H^{(I+)}{}_{1:a}$: trust does not have a positive influence on cooperation

Let us consider a few examples. $H^{(I+)}{}_{0}$: trust does have a positive influence on cooperation; $H^{(I)}{}_{0}$: equilibrium is a mediating variable between trust and cooperation. A standard punch line must be used when the researcher evaluates his hypothesis in the end, after he clearly states his two options:

- $H_{1:0}$: residuals do follow a normal law.[2]
- $H_{1:1}$: residuals do not follow a normal law.

Using an image can help (see Fig. 8.6).

And then, once the researcher has the answer (Fig. 8.7).

[2] Some scholars reverse the order as follows: $H_{1:0}$: residuals do not follow a normal law. However, the rule is that the null hypothesis is an equality. It will take the form of =, ≤, or ≥.

Note that technically, a hypothesis is never actually confirmed. The best one can do is assume it is likely valid. Once the researcher assesses his model like this, testing all possible links between constructs, he can then clean up the final model by keeping only the links that are of value from the point of view of data percolation (from all five angles of analysis). What is neat about quantitative analysis is, as mentioned earlier, that it can help reinforce the model by better explaining the nature and the strength of the links between the variables.

8.5 The Questionnaire

The researcher prepares the quantitative questionnaire by taking into consideration the statistical measures that he wants to take with respect to the model.

Even though a 7-point Likert scale is recommended, they do not apply to ALL questions. For example, socio-demographic questions are not answered with such scales. It is most important that the researcher determines what exactly he is trying to test with his hypotheses (the type of research and subsequently, the links or connections that exist between the variables) and then that he determines what type of data he is going to seek. As many researchers know from reading numerous books on statistics, data come in different forms: nominal, ordinal, ratio, or continuous. What the researcher wants to keep in mind is that he must identify the kind of data he has, most particularly because not all statistical techniques apply to all types of data. Typically, for example, a regression applies to continuous data (with occasionally some binary 0–1 variables). The researcher should also determine if his data are paired or independent. This too will determine what statistical analyses he can perform.

According to the data percolation methodology, the researcher must ask himself four basic questions before finalizing his questionnaire, collecting his data, and analyzing it:

1. "What exactly I am trying to measure?
   a. Is there a link between two constructs?
   b. What is the strength of that link?
   c. What is the nature of that link (e.g. mediation, causal, etc.)?
   d. What is the sign of the link (positive, negative)?
   e. Is there a difference between two groups?
   f. Is there a frequency?
   g. Is there a quantity?
2. What type of data do I have (nominal, ordinal, etc.)?
3. What kind of data do I have (independent? Metric or parametric?)?
4. How many variables do I have?"

The researcher should not produce and confirm hypotheses until he can respond to these questions. For regressions, he also has to determine the normality of the population and of the residuals because regressions are based on a normal law. Many studies forget to test the residuals—this is a mistake.

It is also strongly recommended that the researcher does not initially analyze the data directly by computing it into software such as SPSS or Amos. He should spend a few hours looking at it. He wants to develop a feel for it; he can even do some computation by hand. That technique is part of the data percolation methodology, because it allows the researcher to experience the data rather than plotting it without understanding it in the hope that the software will do the job it is supposed to do. It will, of course, but it will not give the researcher the gut feeling that is so essential to data percolation.

8.6 Distributing the Questionnaire

Sending questionnaires by mail or e-mail can make sense: it is relatively cheap and it avoids contact with the respondent if that is what the researcher is trying to do (in order to pretend to be objective—or be a so-called "positivist"). However, it contains one flaw that should be acknowledged (but that seldom is) in any study that uses this method: the researcher has absolutely no guarantee that the person who claims to have filled out the questionnaire is actually the person who did it, and he has no control over the length of time it took the respondent to do so (it could take 3 days, during which the participant's state of mind certainly changes).

There is another option, which is a basic technique of data percolation methodology. It is called the *live-distribution questionnaire*. In the author's research with seven car dealerships, the author physically went to the waiting rooms where customers were waiting for their cars to be fixed. He would then talk to them and explain he was doing a study (saying he is a student automatically arouses some sympathy) and that he would like to borrow 10 mins of their time (time wasted waiting anyway) to fill out a questionnaire. In 3 weeks, he had more than 200 names. Chances are that questionnaires sent by mail (or even by e-mail) take longer to come back to the researcher.

The live-distribution approach achieves a number of objectives: (1) the researcher is sure the person who fills out the questionnaire is indeed the person who fills out the questionnaire! (2) he can measure the time it takes to fill it out (if customers take too long, he can push a bit by saying he has to go); (3) he can check whether what the respondent answers in the socio-demographic section of the questionnaire corresponds to what he sees (on a few occasions, women who looked well into their sixties responded that they were 40 years old); (4) many times, respondents will come and talk to the researcher after completing the questionnaire. These customers feel obliged or are intrigued and want to express their opinions: this is a gold mine. The quantitative and qualitative data are collected at the same time—data that the researcher can cross-check on the spot! Also, most of the time, grumpy customers will not answer a questionnaire, so the researcher misses an opportunity to have contrasting cases in his data (a fact which may explain why the Customer Satisfaction Index ranks satisfaction high, yet sales remain poor). By seeing other people happily filling out the questionnaire, some (not all) of these unhappy customers

will take it upon themselves to follow the crowd. This method can be used when doing research on the psychological predisposition to buy in a mall, for example. The same applies to playing at the casinos: are people bored at home with the result that they try to step out of solitude by engaging a mechanical relationship with a machine? Or else, do they dream of riches and force themselves to believe that luck is on their side, no matter how poorly they fare in the end when gambling? Or are they already addicted to some games and express their distress by spending money without clear, realistic objectives? Providing the casino owner would be open to the idea of a questionnaire being filled out by the patrons, a live-distribution questionnaire would provide a wealth of information.

8.7 Conclusion

As can be seen, quantitative methods were discussed in a general manner. Errors that are frequently made were discussed; they have no place in the data percolation methodology. For the researcher, not determining the kind of research he wishes to do, not meticulously preparing his questionnaires and his scales and not identifying and formulating his hypotheses and data are not a recipe for rigor.

8.8 A Short Clinical Case

“Some patients express surprise at the fact that I do not take notes while in a therapy session (I take notes immediately after it ends, however). For some patients, it is perceived as an act of benevolence: they feel I pay full attention to their story. For other patients, however, it might be viewed as somewhat troublesome: they might think I do not care about their story. Hence, from this perspective, the fact of not taking notes is a moderating variable.” (Claire Poulin, psychologist, 2014).

8.9 A Few Questions

A few questions
Has the psychometric value of the questions on the questionnaire been reviewed?
Have hypotheses been aligned with the type of research the researcher is aiming to do?
Has the researcher considered sorting his data according to their characteristics (e.g., continuous, independent, etc.)?

8.10 A Few Keywords

Type of connections	Type of arrows
Arrows and types of research structural (*S)* and functional (*F*), (*D*) = (descriptive) binary (*Sb*) continuous (*Sc, F*) influence (*I*)—positive or negative (relational) direct (*I* + or *I*–) indirect … mediator (*I* + or I–)	Indirect … moderator (*I* ±) longitudinal (*T*) (predictive) (*T*) or (t) causal (*C*)—positive or negative (*C* + or *C*–)
Live-distribution questionnaire	A technique belonging to data percolation methodology whereby the researcher distributes the questionnaire in person and tries to collect qualitative information at the same time
Moderating variable	An external factor that influences the strength of the bond between an antecedent and a consequent variable
Mediating variable	A variable positioned as an alternative path to the direct path between an antecedent and a consequent variable

8.11 A Few Tips

The researcher should:

- Read a book on quantitative methodology
- Identify all the parameters of his quantitative research, hypotheses, scales, and data
- Not arbitrarily create questions in a questionnaire
- Not work on the basis of a faulty model