A Polanyian Response to "Psychology's Renaissance"

> Knowledge that we hold to be true and also vital to us, is made light of, because we cannot account for its acceptance in terms of a critical philosophy. We then feel entitled to continue using that knowledge, even while flattering our sense of intellectual superiority by disparaging it. And we actually go on, firmly relying on this despised knowledge to guide and lend meaning to our more exact enquiries, while pretending that these alone come up to our standards of scientific stringency (*PK*, 354).[1]

Michael Polanyi's post-critical philosophy resists the idealization of rule-governed scientific knowledge to the neglect (and derogation) of personal judgments that tacitly direct our choice of hypotheses, relevant evidence, procedures of observation, etc., and it counts as misbegotten the notion that truths and falsehoods can be readily distinguished by strict adherence to methods apart from the personal participation of individual researchers. It is my belief that the renaissance in psychology recently described by Leif Nelson, Joseph Simmons, and Uri Simonsohn[2] expresses a longing for just such a vision of science, and though the vision is tempered by the authors' recognition of its final impossibility, the path it lays out demonstrates and sustains a myopia that Polanyi's epistemology does not accept. But the purpose of this paper is not to find fault with the reforms these authors say characterize psychology's new epoch—they are valuable to the discipline as it currently conceives itself; the purpose is rather to confront the fact that personal judgment cannot be dispensed with in psychology, whatever methodological improvements are made, and to show that when faith in our powers of judgment is restored (an objective that is fully achieved in *Personal Knowledge* and only pointed toward here), possibilities for inquiry open that do not oblige the psychologist to decompose persons finally into dissociated elements, but instead make the latter partial clues to the former, who can be approached only by understanding.

The above will be accomplished by first pinpointing the heart of psychology's renaissance and bringing Polanyi's thought to bear on it. For this purpose, Polanyi's treatment of probability (*PK*, Chapter 2) will prove especially relevant; evaluation of it will show the personal nature of null hypothesis testing and how experience must be atomized for treatment by this method. The question of logical levels in psychology will then arise and compel acknowledgement that I-Thou relations constitute encounters that call for strategies of investigation rather unlike those honored in psychology today. An admittedly "green" sketch of one alternative approach to psychological research and some attendant consequences of pursuing it will then be described in closing that, one hopes, carry sufficient value on the face of things to warrant extended treatment in the future.

Before proceeding along this path, however, a clarifying note is in order. While psychology's renaissance is not confined to any particular subfield of the discipline, it is true that social psychology, that specialty concerned with questions of individual motivation, action and thought in relation to others, has experienced sufficient trouble of late to warrant particular attention. Indeed, Nelson and colleagues list five "consequential events" (512) that led to psychology's renaissance, and all of them at least implicate social psychology (513-514). It is for this reason that my paper will emphasize this domain of psychological science over others and employ the generic label "psychology" throughout to refer to it.

I. What is "Psychology's Renaissance"?

Psychology's renaissance, according to Nelson and his coauthors, is an awakening to dubious yet longstanding research practices in the field that elevate the number of false positives (erroneous claims of statistically significant results) in published studies to unsettling levels (517). What fundamentally is in question here is the validity of probabilities psychologists rely on for hypothesis testing—judging, for instance, whether the numerical difference between two conditions of an experiment is the consequence of random fluctuations attributable to sampling procedures or the experimental treatment—and what is sought are practices that more effectively contain the risk of human error in psychological research and, in turn, offer findings that warrant confident allegiance by all who attend to them. The authors express a measured urgency when they say, "[w]hen false positives populate our literatures, we can no longer distinguish between what is true and what is false, undermining the very goal of science" (520). It is clear, then, that psychology's claim to legitimacy is at stake in a most fundamental way. What is the greatest source of threat? It is an activity the authors call *p*-hacking. (The *p* refers to probabilities consulted for decision making in hypothesis testing.)

When confronted with a dataset that took considerable time and many resources to compile, it is not unusual for researchers to evaluate a hypothesis by taking multiple passes at their data, each time modifying, among other things, the measures they analyze, the observations they include versus exclude, and the statistical controls they employ to adjust for potential contaminating variables. The deleterious effects of such practices often go unappreciated even by conscientious researchers, and they were demonstrated powerfully by Simmons, Simonsohn, and Nelson in a 2011 publication.[3] There the authors relied on computer simulations to test the presence of a difference between two randomly selected samples known to come from one-and-the-same population. They examined 15,000 simulations under different conditions of *p*-hacking (e.g., evaluating two measures independently and then combining them for a third test), and on every occasion a test was statistically significant, they documented the result. Based on the criteria psychologists follow in hypothesis testing, we would expect Simmons and colleagues to have falsely detected a difference between the samples about 5% of the time. But *p*-hacking the analyses in one way or another inflated this percentage considerably (7.7% to 12.6%), and when the researchers combined several distinct forms of *p*-hacking (a not infrequent practice in data analysis), the percentage of false positives soared to 60.7%![4] This is why *p*-hacking is counted by Nelson and colleagues as "a first-order problem for the validity of psychological research" (514) and "arguably the biggest threat to the validity of published research" (525).

Psychology's renaissance is not only concerned with *p*-hacking. There are other contributing influences to false-positives including fraud (526) and unintentional oversights in data coding (525). Concerted efforts are being made to correct these problems through increased accountability. Some researchers are making their data publically available, and journal editors are beginning to require full disclosure of study materials, procedures, and analyses from submitting authors (518); others are adopting the practice of pre-registration (519). Here all study details are determined as completely as possible in advance of data collection and recorded as public proof of researchers' intentions, thereby discouraging them from parading free explorations of data as confirmatory tests of a priori hypotheses.

Let it be said again: the concerns and reforms of psychology's renaissance are not trivial. A Polanyian perspective does not change this. No doubt, pursuing truth in community requires keeping that community's house in order, but it is nevertheless reasonable to ask whether the energy behind psychology's renaissance flows from an epistemology that is objectionable from a post-critical perspective. Does the renaissance seek impersonal knowledge, knowledge that is strictly formalized and free from reliance on human judgment to discern reality? The ambition sounds outlandish framed in this way, and its exponents certainly do not claim this as their objective. They recognize the inevitability of human participation in science and the complications that follow from this; they see the need for careful reflection on the use of statistics in research and oppose the mindless pursuit of small probabilities that are publishable (529). Their entire paper, in fact, is predicated on the appreciation that science requires good judgment, and while this gives the appearance that our question should be answered in the negative, it neglects what appears, at least to this author, to be the more overarching message: Failures in psychological science are primarily attributable to human error, ignorance, or caprice, and successes are realized when researchers commit to getting the science right and conform themselves to proper methods and practices.[5] It is a message that echoes what Wolfgang Kohler judged (with some disfavor) to be American psychology's position in the middle of last century: "Our main obligation as scientists is that of avoiding mistakes" (729).[6]

And latent affinities for automation are not impossible to detect in the paper. They come through, for instance, in a curious statement concerning the evaluation of failed replications in psychology. Nelson and his coauthors say, "Just as it is impossible to bathe in the same river twice, it is impossible to run the same study twice." They call this an "*unfortunate fact*" (emphasis added; 520). Assuming these words were chosen conscientiously by the authors, they warrant notice. The inability to conduct exactly the same study twice could only be called "unfortunate" if there was an alternative that is conceived of as ideal and is wished for. What might it be? A world perhaps where a long line of identical studies succeed in revealing the exact conditions under which a range of well-measured outcomes return their predicted values. Such an arrangement would constitute the reliable advance of knowledge that finally stands up on its own and speaks for itself, untarnished by human wiles.[7] Knowing this is not our reality, of course, we are left to acknowledge our state of affairs as an "unfortunate fact" and reduce our expectations to an unattainable *ideal* of objective knowledge that survives hard times better (by being unattainable) and encourages our best efforts to come near it with proper techniques—statistical analyses and probabilities among them. But inasmuch as "[m]an has a pathetic need for rest and safety," (105)[8] we may be tempted again to confuse the ideal with real possibility and to treat statistics and probabilities as the avenue of rescue. Polanyi however shows that even these are reflections of our own ingenuity and reliance on likeminded explorers in pursuing truth, and since the question of probabilities in hypothesis testing is so important to psychology's renaissance, considering his evaluation of the subject more carefully is recommended.

II. Polanyi on Probability and Order

Polanyi reminds us that "[p]robability statements can never be strictly contradicted by experience" (*PK*, 21). Given a bag labeled to contain 95 white marbles and 5 black, our trust in the accuracy of the label and the theoretical probabilities calculated from it, is not fundamentally

challenged when we shake the bag thoroughly and then draw a black marble from it on a single occasion (*PK*, 23). However much we would be surprised by the occurrence, the numerical probability does not rule it out as impossible. Extending the number of trials offers a better test of the bag's purported contents by bringing the force of accumulated evidence to bear on the question (ibid.). But even here, Polanyi tells us that the label and associated probability statement may only be "controverted" (i.e., contested), "not contradicted" (*PK*, 22). Should I draw a black marble from the bag five times on 15 attempts (where the selected marble is returned to the bag on each occasion and the contents shaken), it can be demonstrated that the theoretical probability of this event is *astonishingly* small (it should occur only once in approximately 1,400 replications of the 15 trials!), but not strictly impossible.[9] However preposterous it may seem to raise the question, we can still ask in this case, "Ought we to count the label as mistaken?" Evaluating the observation against the theoretical probability does not force our hand toward rejecting the label; it leaves the matter to us, the inquirers. But the degree of surprise we feel in relation to the observation and the theoretical probability we calculate as a numerical analog of this feeling offers guidance. "I shall be surprised," Polanyi says, "to a degree corresponding to the reciprocal of this numerical probability. Such is my participation in the event to which a probability statement refers, *and this I regard as the proper meaning of its probability*" (emphasis added; *PK*, 22). And yet, in acknowledging this, Polanyi does not "ascribe subjective meaning to the probability of an event," but "universal validity" (ibid.).

How can this be? A clue to an answer may be found in Polanyi's parenthetical confession, "I am prepared to follow [Sir Ronald Fisher]," the 20th century innovator of null hypothesis testing (*PK*, 23). Fisher's method of discerning when an observed result should be counted as genuine or attributable to chance variation is discussed by Polanyi in relation to an experiment of Charles Darwin that compared the heights of self- and cross-fertilized stalks of wheat (*PK*, 22). The average height difference Darwin observed (the cross-fertilized plants were, on average, 20.93 eighths of an inch taller than those that were self-fertilized) was judged by Fisher to be genuine because its probability of occurrence in a distribution assuming no difference at all fell below 5% (ibid.). It is a question, then, of when we should be struck by an experimental result, and Fisher answers by recommending that when the probability of observed discrepancies are found in a theoretical distribution of mean differences centered on zero (no effect) to be less than .05, we should take notice. Polanyi judges this strategy to put his feelings of surprise in sharper contact with reality, and in so doing he willingly submits to the instruction Fisher provides. His surprise and his interpretation of the probabilities that reflect it, therefore, is not untutored; it is conditioned by Fisher's guidance and also ratified by others who accept Fisher's leadership in this place. This responsible act of following is Polanyi's affiliation with a tradition of inquiry that he believes has commerce with the truth, and in his affiliation, which involves submission to the standards of the guild, his surprise is elevated form the subjective to a personal clue to genuine discovery.

It is in recognition of the personal nature of probability statements that Polanyi goes some length to argue against Frege's treatment of language in which declaratory sentences (e.g., "It is raining") are distinguished from statements of assertion ("It is asserted that"). "If language is to denote speech," Polanyi says, "it must reflect the fact that we never say anything that has not a definite impassioned quality" (*PK*, 27). Impersonal assertions of the form, "It is asserted that" are, therefore, no good, and neither are declaratory statements unattached to any human knower

("[It is] no better than an unsigned check; just paper and ink without power and meaning." [*PK*, 29]). Both, for Polanyi, are incomplete symbols—akin, he says, to "a solitary question mark or exclamation mark," (*PK*, 27). Statements of probability are no different. For a person to say in good conscience, "*I believe* [an assertion] the probability of drawing a black marble 5 times out of 15 trials from a bag of 95 white marbles and 5 black is $7\ X\ 10^{-4}$ [a probability statement]" is for that person to "set his seal" (*PK*, 29) to the statement; he acts as a responsible human being who has elected to uphold a particular human tradition he believes to have contact with reality.

This brief review of Polanyi's treatment of probability leads us to conclude that even granting the methodological improvements of psychology's renaissance, great acts of personal commitment remain necessary to make sense of the probabilities it seeks to purify. We must believe them to be true, and once believing, we must decide what to think by their light in a fellowship of like-minded explorers. This entails, as already suggested, receptivity to education by those we accredit as having authority in this domain.

III. Null Hypothesis Testing and Random Sampling

But there remains another feature of probabilities in psychological research that Nelson, Simmons, and Simohnson appear to pass over completely. It lies slightly upstream from *p*-hacking, but is no less relevant to the veracity of hypothesis testing in psychology. Indeed, as best I can judge, for *p*-hacking to pose a danger requires that this earlier problem be resolved, and because it provides another occasion for glimpsing the influence of personal judgment in the assessment of probabilities, it seems fitting to discuss here. Doing so requires that we move further in the direction of null hypothesis testing. A classic 1974 experiment from the intersection of social and cognitive psychology will serve as an example.[10]

Elizabeth Loftus and John Palmer showed participants footage of a car collision; a subset of them were asked following the film, "About how fast were the cars going when they *smashed* into each other?" Another group received the same question, but the word "hit" was substituted for "smashed." The critical question--pointing to the importance of question wording in interrogations and interviews--was whether speed estimates reported by participants differed across the two conditions.

In classic null hypothesis testing, this question is answered by appeal to a theoretical distribution of all possible mean differences (e.g., in speed estimations across conditions) formed by random sampling from a well-defined population in which *no difference* exists. The probability of the *observed* mean difference (or one more extreme) is evaluated within this theoretical (or null) distribution. When the null distribution is rejected as the source of the observed data, this means that the results are unlikely attributable to chance variation resulting from the probability-based sampling procedure the observations were obtained by. In other words, the results are not an artifact of random sampling, but are genuine. Loftus and Palmer do precisely this when they report, "The mean estimate of speed for subjects interrogated with *smashed* was 10.46 mph; with *hit* the estimate was 8.00 mph. These means are significantly different, $t(98) = 2.00$, $p < .05$" (587). A translation of this statement might read, "The probability of observing a mean difference of 2.46 mph (or larger) in a distribution of differences centered on zero and formed by random sampling from a well-defined population is less than .05."

Loftus and Palmer, however, did not randomly sample form a well-defined population (participants were 150 available college student), and it is widely appreciated that psychologists rarely engage in this laborious and expensive enterprise. The interpretation of *p*-values under such circumstances is unclear. In his critically acclaimed textbook that saw five editions over 40 years, educational psychologist William L. Hays warned students in this way: "Inferential methods apply to probability samples, drawn either by simple random sampling or in accordance with some other probability structure. *There is no guarantee of their validity in other circumstances*" (emphasis added; 227).[11] That psychologists report *p*-values despite this, and that psychology's renaissance is so deeply concerned about their purity even in the absence of probability-based sampling procedures, is a curiosity. One is left to conclude that psychologists dislike the play of personal judgment in their evaluation of evidence any place they find it, except when it is convenient to the work. Treating as inconsequential differences between their available samples and those they might hypothetically have obtained through random sampling seems an example of this. But the proper response to this state of affairs is not, for Polanyi, stricter methods, per se, but to trust the powers of judgment that led us to the method in the first place and then continue to direct our steps when the explicit guidance supplied by the chosen method fails or falls silent. Indeed, I take this to be exactly what psychologists do when they (responsibly) decide to interpret the probabilities of their research under circumstances where the requisite assumption of random sampling is not satisfied.

But there is another option that accrediting our powers of judgment raises. Hays says, "Unless the assumption of random sampling is at least reasonable, the probability results of inferential methods … *might as well be omitted*" (emphasis added; ibid.). If it is possible to responsibly wield null hypothesis testing in the absence of strict random sampling, might it also be possible for a psychologist to responsibly leave this method behind altogether under appropriate circumstances? The force of this question gains power when we appreciate a further observation Polanyi makes in relation to probability and order: he says that it is "only in view of … orderliness that the question [can] be asked at all whether the orderliness [detected] was accidental or not" (*PK*, 34). When we evaluate the probability of a particular observation, we do so precisely because the outcome in question has struck us, standing out to our eyes against a background of fluctuations that, according to the conception of "events governed by chance," only could have produced what we see "by coincidence" (*PK*, 36). Null hypothesis testing is the method by which we evaluate the observation's likelihood against chance, but the fact of the observation striking us in the first place is a testimony to the trust we place in ourselves—as, it should be added, is the decision to dismiss the observation out-of-hand and not make the test at all. So whether we make the test or we do not, it seems from a Polanyian perspective that we already accredit our capacity for distinguishing order from randomness. Turning to Fisher's strategy in light of this offers a valuable tutelage in handling our surprise through "a partial formalization of a personal act" (*PK*, 30), and we accept the instruction in the proper places, but it does not dictate belief by distinguishing impersonally and precisely for us order from randomness. This we must finally judge for ourselves; there has never been an alternative.

IV. What Follows from Granting the Reality of Logical Levels?

Reference has been made now on more than one occasion to accepting Fisher's guidance in the proper places. What are those places? At present I have only a tentative answer. It arises from the supposition that Polanyi would always invite us to consider the question of "the proper place" and to responsibly seek answers in a community of others similarly committed to knowing truth. Based on the examples provided in Polanyi's chapters on probability and order, we might infer that the proper places for such analyses are those that resemble to us questions concerning marbles in bags or the heights of differently fertilized plants. And it is worth appreciating that such segmented and simple phenomenon are arrived at in psychology only through *analysis* where wholes are decomposed into simple, dissociated units of observation treatable by measurement and experiment.[12] The brief example offered above in Loftus and Palmer's research pictures this; individuals' recollections were confined to what they could report in a "Yes/No" format. Such examinations are not unimportant[13], but there is obviously much more to the story, and this is granted when the individual psychologists investigate is properly appreciated as a Thou rather than an It, a conclusion that readily follows from Polanyi's discussion of logical levels.

According to Polanyi, a two-tiered logical structure holds in the scientist's investigation of inanimate matter: there is the object itself (the first logical level) and the scientist's knowledge of the object as a sample of quartz, silt or clay, etc. (the second logical level). This is a "*knowledge of things*" (*PK*, 344). It is distinct from our "*reflections on our knowledge of things*" (e.g., "the logic and epistemology of science"), which constitutes a third logical level (ibid.). This third level surfaces anytime we think about our thoughts about things *and* when we study living organisms in light of what *they* know—a rat's mental map of a maze, for instance (*TSOM*, 76).[14] This situation defies the two-tiered logical structure of physical science by involving reflection on *another being's* knowledge, and it follows from Polanyi's commitment to personal knowledge and the process of evolution he believes gave rise to the mental powers he recognizes in himself that he accredits to lower organisms primitive manifestations of the same. Studying the knowledge of these organisms thus entails comprehending what they know, discerning what they intend to do, rendering a judgment about the efficacy of their knowledge, and, perhaps, endowing them with new knowledge through education. It *does not* entail treating them as lifeless objects or mindless machines, as befits firewood and toaster ovens (*PK*, 344-345). This is a crucial shift for Polanyi, and it grows in importance when inter-human relationships are considered. Here "[t]he I-It situation" characterizing the study of inanimate matter "transform[s] into an I-Thou relation" imbued with an altogether distinct degree of "[m]utuality" that is only prefigured in the study of lower animals (*PK*, 346; *TSOM*, 33). The investigator and the subject of investigation now have the potential for comparable self-understanding and contact with reality. The investigator's knowledge of the subject in this case "has lost the character of an observation and has become an *encounter* instead" (emphasis added; *TSOM*, 95). This does not mean that the depth of understanding or contact held by the two are always on a par, but it does preserve the potential for this to be so, and also the potential for the subject to surpass the investigator in these areas.

Treating individuals atomistically without any or with only the faintest view toward the whole of their mental existence neglects this "mutuality." This neglect carries the practical advantage of offering simple facts that can be translated into averages or percentages, of course, but looked at honestly, such quantities are only *clues to still further investigation*. They lead on to vital shades

of meaning in the individual and the community he inhabits. Rather than ends in themselves, they are the earliest beginning of comprehension, and it seems, at least to this author, that a psychology which purports to speak authoritatively about the experiences of persons in their social surroundings could do better. Two examples will serve as illustrations.

Questionnaire responses in which individuals are required to report their level of agreement (1 = *Strongly Disagree*, 7 = *Strongly Agree*) with a series of similarly themed statements are widely employed by psychologists. For instance, a popular instrument for measuring self-esteem includes such statements as, "I wish I could have more respect for myself," and "At times I think I am no good at all."[15] Now what would it mean for an individual to agree strongly with such statements (indicating low self-esteem)? It could mean he is haunted by moral failings he has not made amends for. His conscience plagues him each time he remembers his crimes, and he remembers often. For another it could mean that her professional dreams of becoming a physician have been dashed by a failed entrance exam; she is endeavoring now to remake herself and she oscillates daily between hope and doubt about the final success of her efforts. For still another it could mean "the knife is always at the wrist." This person feels his soul to be as empty as the cosmos. Lost and alone, he would rather not exist, but he lacks the constitution to finally end things; he is miserably stuck with himself. But none of this precious information is available in ratings of agreement. It is sacrificed for simple scale averages because nuance cannot be treated by statistics. It is in this sense that such ratings offer clues. They have the potential to point, to suggest, to indicate, but less so to reveal or make known.

The second example comes from Stanley Milgram's famous obedience studies. His participants were fooled into believing they were administering increasingly painful shocks to another innocent person. It is well known that his procedure evoked intense emotional responses from participants. According to one observer, a man who proceeded to administer the strongest shock possible (450 volts) was "rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point he pushed his fist into his forehead and muttered: 'Oh God, let's stop it'" (377).[16] Now is this behavior best interpreted as a simple instance of obedience (which it was)? His acts seem overshadowed by the powerful external signs he presented of internal distress. Suppose, for instance, that years ago the man was the punching bag for an abusive, alcoholic father. The stone-faced appearance of the experimenter churned up memories from the depths, leading him momentarily to revert to the gutted personality of his pre-pubescent self and relive the submissiveness that saved him from his father's violence. What richness such insight would add to the opaque label "obedient," but it is avoided. Why?

The psychologist avoids treating his quantities as clues to wholes—as means to more important ends—not just because he wants to use statistics (a misbegotten mark of scientific legitimacy in the human sphere), but also from his fear that bias pervades his personal judgment and the corresponding worry that allowing himself to encounter comprehensive entities who are finally irreducible to their measurable parts would require him to rely on it in research. A question psychologists should grant far greater importance than they do is, "What are the *meanings* of this or that individual's utterances? And how are the expressions of several or many members of a community properly brought together into a reflection of their shared experiences?" For such questions to be accepted in psychology as worthy of investigation, the answers that follow would

have to be trusted as more than merely interpretive or subjective. A full defense of why such trust is warranted would require a recapitulation of *Personal Knowledge* in its entirety, and this, of course, is beyond the scope of the paper. But I do hope the above remarks show at least something of the operation of personal judgment where we are unlikely to appreciate it (e.g., hypothesis testing), and that they help smooth the way for considering what follows.

To be clear, however, it would be mistaken to conclude that what has been argued up to this point is simply that one mode of inquiry can responsibly be substituted for another once the personal coefficient of our knowledge is accepted. This is fundamental, but it does not stop there. In light of the "I-Thou" relation discussed above, *such substitution is called for* when the goal is genuine understanding of others' experiences.[17] Accepting this goal as the psychologists' highest obligation, as I do, certain consequences follow, including the repositioning of results acquired by analysis (experimentation, quantification, etc.) to lower rungs of importance as clues (among others) to wholes and the elevation of insights gained by broader inquiries (like interviews and active listening discussed below) to higher ones. This is what follows from granting the reality of logical levels in psychology.

IV. Understanding, Looking and Listening

It is now necessary to bring clarity to the notion of understanding mentioned at the beginning of this paper and again in the preceding section. Once done, a possible means of pursuing it in psychological research will be discussed. In the opening chapter of *TSOM*, Polanyi describes understanding as "the word which covers" such tacit acts as "see[ing] things in one way rather than another" and learning "our way about a neighbourhood." All of these acts consist, he says, "in comprehending experience, i.e. in making sense of it" so as to achieve "intellectual control over it" (*TSOM*, 19-20). Inquiring into the meaning of understanding for Polanyi is inquiring into tacit knowing, and this he aptly depicts by appealing to the way in which men learn to navigate unfamiliar territory (*TSOM*, 14-15). Polanyi says that "[i]n the absence of linguistic clues man sees things, hears things, feels things, moves about, explores his surroundings and gets to know his way about, much as animals do" (*TSOM*, 14). It is through the communications of his bodily instruments--eyes, ears, skin and fingertips--and his tacit integrative capacities that the layout of new environments rise to coherence in his mind. A competent tracker may represent his knowledge of a vast wilderness on paper for posterity and may enrich his own competency by referencing such tools, but neither the initial convergence of experience that made his map-making possible nor exhaustive instructions to future users of how to bring the map properly to bear on experience can be articulated. All the pointing and explaining and correcting in the world is for naught when understanding is the goal. The tracker and those he wishes to instruct must rely on powers of synthesis that defy specification in order to become connoisseurs of a land's topography.

An example closer to the problems of psychology may bring further clarity. "We know a face," Polanyi says, "without being able to tell, except quite vaguely, by what particulars we recognize it. And this is also how the mind of man is known" (*TSOM*, 33). Confronting a stranger we wish to know is not unlike the situation faced by a tracker in novel surroundings. We are in uncharted territory, our senses wide open to the stranger, his manner of dress, countenance, tone of voice, and gesticulations. His "mind can be known [to us] *only comprehensively, by dwelling within the*

unspecifiable particulars of its external manifestations" (ibid.). These clues (or "unspecifiable particulars") point to his inner reality, and they call more for lingering over than for measuring, just as one does when laboring to comprehend a writer whose expressed thought invites by suggestion and hints the pursuit of an undisclosed meaning. It calls to mind the preparations an actor makes for a monologue. By sinking into the lines written for the role, the actor finds fertile soil for releasing what eventually extends into a dynamic and integrated picture of the emotion and thought of the character. Such integrated pictures appear in flashes of coherence like the perceptual totalities the Gestalt psychologists discovered, and this, fundamentally, is what Polanyi means by understanding. He calls it the "grasping of disjointed parts into a comprehensive whole," a pre-articulate capacity for integration we share with animals (*TSOM*, 30).

Wilhelm Dilthey's definition of understanding—though not his grasp of its relevance (expressed through indwelling) to both the sciences and humanities[18]—harmonizes with Polanyi's. "We must … call Understanding," he says, "that process by which we intuit, behind the sign given to our senses, that psychic reality of which it is the expression. Such understanding ranges from the comprehension of the babblings of children to *Hamlet* and the *Critique of Pure Reason*" (232).[19] Because of Dilthey's special concern with understanding in historical research, he goes on to add, "If, for instance, I wish to understand Leonardo, my interpretation of his actions, paintings, sketches and writing works together as a single homogenous and unified process" (ibid.). But what of persons now living, those who are not the concern of history, but of psychology? And if null hypothesis testing only offers "disjointed parts" to investigators that must be viewed with other clues as pointers to wholes, the question arises: how is research to proceed? The answer I can offer here, while evidently inchoate, nevertheless carries the benefit of being straightforward and unassuming: *We might begin with looking and with listening*.

Pursuing another's inner experience of himself and the world by patient and deliberate attention to his communications requires conditions in which these communications may be freely expressed. A student who is apprehensive about her comprehension of a lecture does not volunteer this to her professor if she fears his evaluation. Far too much is at stake, and in her silence the professor himself occupies a position of ignorance. He is unaware of the contours of her knowledge and so cannot find expressions to raise the prospect of her learning. Rectifying this situation requires that he alter its conditions by becoming a student of *her* thought. A form of submission ensues for a time in which the teacher suspends the objective of training, correcting, and evaluating so that the goal of seeing with the pupil assumes ascendancy. Within certain forms of psychotherapy[20] too this asymmetric relationship holds, but here it is not only facts about the self that the other must feel free to convey, but also, and vitally, affect. Seeing with the other now calls for following the course of her emotional currents, and for both the facts and the feeling to flow freely in communication, the therapeutic relationship must be accompanied by an atmosphere of genuine interest and empathy sustained by the psychologist. Such conditions carry a definite power to release others to free and capacious communication, and it is within this bounty of clues that psychologists must dwell for the inner lives of others to take shape for them.[21]

One wonders whether looking and listening of this kind, carefully adjusted and supplemented by artifacts similar to those relevant for understanding Leonardo, have application beyond

investigations of psychotherapy to the problems of psychology more generally. How might certain mainstays of psychology—e.g., social influence, attributions, prejudice, aggression—be refashioned by this approach, and might something akin to it be better adapted to the problems of individual differences and cross-cultural research by being situated in a particular time and place with definite persons who can be seen and heard? Furthermore, and importantly, might such an approach be more consistent with the logical levels and I-Thou relations Polanyi emphasizes? All of this requires extended reflection, but granting for the moment that such questions can be answered in the affirmative, what in very general terms would follow?

It seems that reliance on experimental methods and inferential statistics would diminish to make room for encounters with the free communications of others. This shift would, in turn, require psychology students' training to broaden. The pattern of looking and listening sketched above evidently resembles the case studies of clinical psychology, the hermeneutical approaches described by others[22], and when extended to communities or groups, also the ethnographic and participant-observer techniques of anthropologists and journalists. A reoriented psychology would draw upon the insights of these other disciplines and perhaps intermingle with them; and as Sigmund Koch's piercing observation at the turn of the century makes clear, a reoriented psychology would also have much to gain from the humanities.[23] Just how much is uncertain, but if training in great literature only served to burden psychologists again with the question of man's totality, an important service would be done.

V. Conclusion

Is all of this worrisome? Perhaps to a point. But it is worth recalling that Kurt Lewin himself, the accredited founder of experimental social psychology, emphasized individuals' inner experiences (what he called "the field" or "life space" of the person) as the sphere of inquiry for psychologists. "*Objectivity* in psychology," he said, "demands representing the field correctly as it exists for the individual in question at that particular time. For this field [the individual's] friendships, conscious and 'unconscious' goals, dreams, ideals, and fears are … essential" (338).[24] And, lending credence to Koch's position above, he even held out Dostoevsky's work as exemplary, saying that the knowledge he had of his characters is the kind psychologists must strive after with others (13).[25]

It is also worth considering a proposal Polanyi made in *Science*[26] and later shared with Carl Rogers in a televised conversation. "If we could only get away from [the word 'scientific' for ten years]" he said, "we would see so many possibilities of appreciating knowledge—of appreciating views and explorations" that we might rightly call instead "penetrating, revealing, sensitive, [and] true … It is quite an obvious way of describing them" (159).[27] And regarding Rogers' experience inside therapy (he felt conflicted about its scientific legitimacy), Polanyi expressed "complete confidence in the value of such a pursuit" to the advancement of truth. Whether the experience was quantitative or qualitative, "scientific" or otherwise, did not trouble Polanyi: "It seems to me," he said, "not a substantial question" (ibid.,175). In speaking these words, Polanyi offered to Rogers what the latter observed in therapy with clients: *freedom from tension*.[28] Were such freedom to permeate psychology today, what marvelous possibilities would unfold, and what truly vibrant inquiries might be permitted to attend the field's (latest) rebirth.

Endnotes

[1] Polanyi, M. 1958. *Personal Knowledge* (Chicago: The University of Chicago Press). Parenthetical references in the paper using the abbreviation "*PK*" refer to this work.
[2] Nelson, L., Simmons, J., Simohnson, U. 2018. Psychology's Renaissance. *Annual Review of Psychology*, 69:511–534.
[3] Simmons, J., Nelson, D., Simonsohn, U. 2011. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22:1359-1366.
[4] See Endnote 3, Table 1 on page 1361.
[5] Nelson and his coauthors, in fact, reference the Internal Revenue Service's auditing practices as a model for checking the validity of psychologist's published work. "We think that researchers should … behave as if they will be audited. For example, journals could require authors to provide information on exactly when (i.e., specific dates and times), exactly where, and by whom the data were collected. Journals could then do the routine fact checking that newspapers do" (527).
[6] Kohler, W. 1959. Gestalt Psychology Today. *American Psychologist*, 14:727-734.
[7] Recall one of the opening lines from *The Study of Man* (see Endnote 14 for full reference): "Man must try forever to discover knowledge that will stand up by itself, objectively, but the moment he reflects on his own knowledge he catches himself red-handed in the act of upholding his knowledge" (11-12).
[8] Kohler, W. 1949 (1992 reissue of 1970 paperback). *Gestalt Psychology* (New York: Liveright Publishing Corporation).
[9] $\binom{N}{r}p^r q^{N-r} = \frac{N!}{r!(N!-r!)}p^r q^{N-r}$, where N = number of trials, r = number of successes, p = probability of success, and q = probability of failure.
[10] Loftus, E., Palmer, J. 1974. Reconstruction of Automobile Destruction: An Example of the Interaction Between Language and Memory. *Journal of Verbal Learning and Verbal Behavior*, 13: 484-489.
[11] Hays, W. 1994. *Statistics*, 5th Ed. (Belmont, CA: Wadsworth Group/Thompson Learning). I am aware of two reviews of the first and second editions of Hays' text. They were published in *Educational and Psychological Measurement* (vol. 24 [1964] and vol. 34 [1974]), and high praises are evident in both. Also, it is important to appreciate that Hays' statement about random sampling should not be taken as a comment on generalizability (i.e., the degree to which conclusions drawn from an observed sample can be extended to a population). The issue here is whether the observed result should be counted as chance variation resulting from a probability-based sampling procedure or as a genuine effect; it is, of course, related to generalizability (random samples are taken to be unbiased representations of populations), but not strictly the same. They concern different implications of a common process.
[12] It is striking how similar this position is to that of Wolfgang Kohler who objected mightily to introspectionists' strategy for studying visual perception: "The psychologist … confronted with a complex field of vision … feels naturally inclined to analyze this field into smaller and simpler entities whose properties he may study with more ease and with more hope of clear results than an immediate consideration of the whole field would yield. Generally he does not ask himself what this procedure purports and if, perhaps, the term analysis is rather ambiguous. He simply

analyzes down to very small parts … dissolving [the field] theoretically and arbitrarily into minute local things which nobody ever sees." (*Some Tasks for Gestalt Psychology*, 147).
[13] "[Destructive analysis] is indispensable, and without its constant exercise no scientist or technician could keep a steady course among the many spurious observation which he has to set aside unexplained every day" (*PK*, 51). But to this observation Polanyi joins a warning: it could lead to "explaining away quite genuine practices or experiences" (ibid.).
[14] Polanyi, M. 1959. *The Study of Man* (Chicago: The University of Chicago Press). Parenthetical references in the paper using the abbreviation "*TSOM*" refer to this work.
[15] Rosenberg, M. 1965. *Society and The Adolescent Self-Image* (Princeton, NJ: Princeton University Press).
[16] Milgram, S. 1963. Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology*, 67: 371-378.
[17] In *The Psychology of Science: A Reconnaissance* (1966) Abraham Maslow asks, "[I]f I want to know a person, what is the best way to go about doing it? … How good for this purpose are the usual procedures of normal physical sciences …? In general my answer is that they are not very good at all. As a matter of fact, they are practically useless if I want not only to know about you but also to understand you" (10).
[18] Polanyi, M. 1962. Tacit Knowing: Its Bearing on Some Problems of Philosophy. *Reviews of Modern Physics*, 34: 601-616. In this article Polanyi says Dilthey "was mistaken in distinguishing indwelling from observation as practiced in the natural sciences. The difference is only a matter of degree: Indwelling is less deep when observing a star than when understanding men or works of art. The theory of tacit knowing establishes a continuous transition from the natural sciences to the study of the humanities. It bridges the gap between the I-It and the I-Thou, by rooting them both in the subject's I-Me awareness of his own body, which represents the highest degree of indwelling" (605).
[19] Dilthey, W. 1900 (Republished 1972). The Rise of Hermeneutics. *New Literary History*, 3:229-244 (Translated by F. Jameson).
[20] Rogers, C. 1947. Some Observations on the Organization of Personality. *American Psychologist*, 2:358-369. Rogers says, "If we take first the tentative proposition that the specific determinant of behavior is the perceptual field of the individual, would this not lead, if regarded as a working hypothesis, to a radically different approach in clinical psychology and personality research? It would seem to mean that … [w]e would try to see with him, rather than to evaluate him. It might mean the minimizing of the elaborate psychometric procedures by which we have endeavored to measure or value the individual from our own frame of reference ... They are not the ways in which the individual experiences himself" (367).
[21] Should this sound similar to the Gestalt phenomena Polanyi discusses, it is no coincidence. The connection was also apparent to Carl Rogers: "Just as, by active concentration, one can suddenly see the diagram in the psychology text as representing a descending rather than an ascending stairway or can perceive two faces instead of a candlestick, so by active effort the counselor can put himself into the client's frame of reference" (32; Rogers, C. 1951 [2015 reprint]. Client-Centered Therapy [London: Robinson]).
[22] Sandage, S., Cook, K., Hill, P., Strawn, B., Reimer, K. 2008. Hermeneutics and Psychology: A Review and Dialectical Model. *Review of General Psychology*, 12: 344-364.
[23] Koch, S. 1993. "Psychology" or "The Psychological Studies." *American Psychologist* 48:902-904. From page 903: "Fields such as sensory, physiological (or broadly neuroscience-oriented) psychology may certainly be seen as solidly within the family of the biological and, in some

reaches, natural sciences. But psychologists must finally accept the circumstance that extensive and important regions of psychological study require modes of inquiry … rather more like those of the humanities than those of the sciences. And among these latter, I would include important sectors of areas traditionally considered fundamental (e.g., perception, motivation, and learning) as well as such obviously more rarified fields as social psychology, psychopathology, aesthetics, and the analysis of creativity."

[24] Cartwright, D. (Ed.) 1951. *Field Theory in Social Science: Selected Theoretical Papers*. Behavior and Development as a Function of the Total Situation (1946) by K. Lewin. (Oxford, England: Harpers).

[25] Lewin, K. 1936. *Principles of Topological Psychology* (New York: McGraw-Hill Book Company). While Lewin praises the achievements of novelists like Dostoevsky, it is nevertheless true that he did so with a view toward methods that this paper calls into question.

[26] Polanyi, M. 1957. Scientific Outlook: Its Sickness and Cure. *Science*, 125:480-484. On page 125 of this article, Polanyi says, "I would suggest that we might begin to remedy this weakness by prohibiting the use of the term scientific in praise of a study of human society, for trial period of, say, 10 years. And in the meantime we should try training ourselves to study human affairs by intense participation in human problems instead of by detachment from them."

[27] Kirschenbaum, H. and Henderson, V. (Eds.) 1989. *Carl Rogers: Dialogues: Conversations with Marin Buber, Paul Tillich, B.F. Skinner, Gregory Bateson, Michael Polanyi, Rollo May, and Others* (Boston: Houghton Mifflin Company).

[28] On page 363 in Some Observations on the Organization of Personality (Endnote 20, above) Rogers says that "as we observe and study the recorded accounts of the conclusion of therapy, it is clear that the most characteristic outcome is not necessarily solution of problems, but a freedom from tension a different feeling about, and perception of, self."