While Waiting for Nature to Take Her Course:  
There’s Nothing So Practical as a Good...Design

Mark P. Zanna*
University of Waterloo

In this essay, I first describe some examples of curiosity-based, basic research inspired by Lewin’s famous dictum that there is nothing so practical as a good...theory. Then, I describe several examples of recent intervention studies that leverage basic research to make the world a better place. Next, in the spirit of Lewin, I suggest that there is nothing so practical as a good...design—and describe “state of the art” quasi-experimental designs to evaluate tobacco control policies and suggest the sort of designs one might employ to evaluate smoking prevention programs implemented before kids have even started smoking. Finally, I conclude by suggesting the need to change our academic culture to encourage more emphasis on mobilizing the knowledge we obtain from our research.

First, let me thank the Kurt Lewin Award Selection Committee for bestowing this great honor on me. Needless to say, it was a big surprise to be considered in the same group as past winners—including several of my graduate school heroes, Fritz Heider, Mort Deutsch, Don Campbell, and Hal Kelly, and several of my close professional friends, Jim Jones, Claude Steele, Faye Crosby, and Mark Snyder. Jim, Claude, Faye, Mark, and I go back so far it would be impolite to say how far!

Second, let me acknowledge my professional debt to Kurt Lewin. Of course, I never met Lewin. I was, however, trained by his direct “descendants.” At Yale,

*Correspondence concerning this article should be addressed to Mark P. Zanna, Department of Psychology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada [e-mail: mzanna@uwaterloo.ca].

Because virtually all my research is a collaborative effort, the research discussed in the present essay was conducted with several graduate students—and a couple of colleagues. Therefore, I need to acknowledge the following individuals, in their “order of appearance,” for their immense contributions to the research: Christian Jordan, Leanne Son Hing, Steve Spencer, Tara MacDonald, Geoff Fong, Sonya Dal Cin, Erin Strahan, Emiko Yoshida, Jennifer Peach, Christine Logel, Danu Stinson, Dave Hammond, and Shelagh Towson. I also thank Gene Borgida, Christine Logel, and Adam Zanna for comments on an earlier draft of this essay. Finally, I thank Steve Spencer for delivering an abridged version of this essay at Society for the Psychological Study of Social Issues’ 8th Biennial Convention in New Orleans—while I was on the “Disabled List.”

© 2011 The Society for the Psychological Study of Social Issues
Chuck Kiesler was my primary PhD advisor; Dick Nisbett was my secondary advisor. Chuck was supervised by Leon Festinger, who in turn, of course, was supervised by Kurt Lewin. Dick was supervised by Stan Schachter, who in turn, was supervised by Festinger. So, I guess I am either Lewin’s great grandson through Kiesler or great, great grandson through Nisbett. At any rate, there is no doubt that my mentors taught me in the Lewinian tradition to conduct high-impact experiments that tested theory relevant to (and often inspired by) social issues. Put simply, when feasible I am inclined to create high-impact psychological states as independent variables and assess consequential behaviors as dependent variables relevant to social issues (see Jordan & Zanna, 2007) for an essay that recommends this strategy for creating “persuasive experiments”).

Third, given this background it should not be surprising to learn that I have been a member of the Society for the Psychological Study of Social Issues for my entire academic career.

**There is Nothing So Practical as a Good... Theory**

*Curiosity-based, Basic Research*

As a “card-carrying” Lewinian, I continue to believe that there is nothing so practical as a good theory. For me, this means (1) doing curiosity-based, basic research that is relevant to the social issues of the day, and (2) as mentioned above, when possible, doing high-impact research with consequential behavioral dependent variables.

For example, recently I have taken advantage of the “loose connection” between explicit and implicit prejudice to identify aversive racists on the left side of the political spectrum by proposing that aversive racists are individuals low on explicit prejudice but high on implicit prejudice. In our first study, we tested this hypothesis in the context of a hypocrisy-induction experiment (Son Hing, Li, & Zanna, 2002). Put simply, when we “rubbed the noses” of our participants in the fact that they do not always practice what they have preached in the context of intergroup relations, aversive racists (so defined), as we expected, felt guilty, and bent over backwards not to discriminate on a consequential budget-reduction ballot (i.e., they voted to minimally reduce the budget of a visible minority club on campus). More recently, we demonstrated that aversive racists only discriminate when there is a subtle excuse to do so (Son Hing, Chung-Yan, Hamilton, & Zanna, 2008). In this follow-up research, we also discovered that not all individuals on the right side of the political spectrum are modern racists. In fact, although modern racists (those high in explicit modern racism and high in implicit prejudice) do discriminate when there is a subtle excuse to do so, principled conservatives (those who score high on the explicit modern racism scale, but low in implicit prejudice) do not—that is, they are indistinguishable from those who are not likely to be
prejudiced (those low on both explicit and implicit prejudice). Thus, although not all individuals on either the left or the right are prejudiced, some clearly are and we can now identify them a priori! So, let us get started on studying the similarities and the differences in prejudice both on the right and the left.

I have also taken advantage of the “loose connection” between explicit and implicit self-esteem to identify defensive individuals by proposing that such individuals are high on explicit self-esteem, but low on implicit self-esteem. In our first studies, we tested this hypothesis by demonstrating that defensive individuals (so defined) are more likely to be defensive than other groups, both dispositionally (i.e., they are the most narcissistic, a personality trait related to defensiveness) and situationally (i.e., they reduce dissonance by rationalizing their decisions the most) (Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003). More recently, we demonstrated that following a threat to their self-esteem defensive individuals are also the most likely to discriminate against an out-group member (Jordan, Spencer, & Zanna, 2005).

**Intervention Studies**

But if there is nothing so practical as a good theory, then we ought to use the theories that curiosity-based, basic research support to design interventions to make the world a better place. I view such intervention studies as both basic and applied. They are basic in the sense that they often test the theory in a new context. And, of course, they are applied in the sense that they are designed to make a difference in the world. Although intervention studies have not always been the focus of my research, they have become more of a focus in recent years.

So, what exactly do I mean by an intervention study? Although the hypocrisy study described above did seem to reduce prejudice for aversive racists, I do not regard it as an intervention study. To be an intervention study, we would have had to create and test the effects of an educational module that induced hypocrisy in students in an actual school context.

Let me now describe a few recent intervention studies that have followed up on curiosity-based, basic research.

**Alcohol Myopia and Condom Use**

For several years, my colleagues and I have tested Claude Steele’s “Alcohol Myopia” Theory (the notion that intoxication reduces cognitive capacity so that individuals are influenced primarily by cues that are momentarily salient) to solve the puzzle of why men who claim to always use condoms do not always do so when intoxicated (Ebel-Lam, MacDonald, Zanna, & Fong, 2009; MacDonald, Fong, Zanna, & Martineau, 2000). In lab studies we have demonstrated that, when impelling cues (such as “she is attractive”) are most salient (as they typically are
in the real world), intoxicated individuals are more likely to intend to have unsafe sex than their sober counterparts, whereas when restraining cues (such as “she might have a STD”) are most salient (as they typically are not in the real world), intoxicated individuals are less likely to intend to have unsafe sex than their sober counterparts. This basic research was exciting for two reasons. First, it supported the alcohol myopia prediction that intoxication does not always lead to risky decisions. Second, it suggested an intervention, namely, to expose intoxicated individuals to restraining cues when they are “in the heat of the moment.” So, that is what we did in a field experiment. Put simply, we added a cue to a standard “safe sex” intervention (in the form of a friendship bracelet) designed to remind participants that they should act on their sober intentions to use condoms. To our amazement, we found that the reminder cue worked and did so to a greater extent when our participants reported they had been drinking (Dal Cin, MacDonald, Fong, Zanna, & Elton-Marshall, 2006).

Norms for Ideal Appearance and Contingencies of Self-Esteem

Colleagues and I have also been interested in testing the notion that the cultural norms for women’s ideal appearance are problematic for women’s satisfaction with their bodies and, ultimately, their eating behavior because the salience of such norms leads women to base their self-worth more strongly on appearance (Strahan, Spencer, & Zanna, 2007). What was interesting (to us) about this research was the fact that we assessed contingencies of self-worth for the first time at the dependent variable (rather than at the independent variable) level. With support for our hypothesis, we decided to conduct an intervention study designed to delegitimize the cultural norms for women’s ideal appearance in a middle school, focusing on the notion that one’s self-worth ought not to be solely contingent on one’s appearance. In one 80-minute session, students discussed the historical changes in “ideal” body types, problem solved solutions to scenarios in which an individual felt pressure to look a certain way, and discussed how unrealistic and unattainable appearance ideals are (e.g., by contrasting the real-life photos of celebrities with their “perfected” media images). In a second 80-minute session, they applied what they had learned in a series of debates and created posters challenging the norms of appearance. Interestingly, we were able to delegitimize the cultural norms for ideal appearance for both girls and boys. And, as a result girls came to feel better about their bodies, mediated by the fact that girls did, indeed, base their self-worth less strongly on appearance (Strahan et al., 2008). Although the delegitimization of cultural norms for ideal appearance did not influence boys’ contingencies of self-worth, the fact that they came to believe the cultural norms of appearance were illegitimate for boys and, especially, for girls was probably critical to the effectiveness of the intervention for the girls. At any rate, the presence versus the
absence of boys ought to be tested in future interventions designed to delegitimize appearance norms for girls.

A “Chilly Climate” for Female Engineering Students

In recent years, perhaps, the line of research about which I am most excited concerns implicit norms. Following Olson and Fazio’s innovation of “personalizing” the traditional Implicit Association Test (IAT) by changing the evaluative category labels from “pleasant” and “unpleasant” to “I like” and “I dislike” (Olson & Fazio, 2004), it was clear to me that the traditional IAT captured two sources of variances: automatic personal evaluative associations (i.e., implicit attitudes) and automatic cultural or normative evaluative associations (i.e., implicit norms). Recognizing that Olson and Fazio had “tweaked” the traditional IAT to capture implicit attitudes, Steve Spencer and I (and, of course, our students), in turn, decided to “tweak” Olson and Fazio’s personal IAT to capture implicit norms by changing the evaluative labels to “most people like” and “most people dislike.” So far, we have demonstrated that although they are often “loosely connected” to implicit attitudes and explicit norms, implicit norms have different/unique causes and consequences (Yoshida, Peach, Zanna, & Spencer, unpublished data). For example, the amount of time Asian students have spent in North America decreased their implicit normative regard, but not their implicit attitudes, toward the elderly. Situations that elicit stereotype threat in women decreased implicit norms toward women but not implicit attitudes. Hearing an audience laugh at a racist joke about people from the Middle East (vs. hearing the same joke with no laughter) made implicit norms toward people from the Middle East more negative but did not influence implicit attitudes—and, as a consequence, resulted in participants discriminating more against Muslims (i.e., by voting to cut more funds for a Muslim student organization on campus). Implicit norms toward Blacks predicted response times on a shooter bias task (such that those with more negative implicit norms were faster to shoot a Black target holding a gun and slower not to shoot a Black target holding a cell phone, cf., Correll, Park, Judd, & Winterbrink, 2002), whereas implicit attitudes, explicit attitudes, and explicit norms did not (Yoshida, Peach, Spencer, & Zanna, 2008).

Most important for the present essay, we also successfully “bottled” the chilly climate for female engineering students by assessing implicit normative evaluations of women in engineering at the beginning and end of their first year. We found not only that these implicit norms became more negative over the year—that is, students came to automatically associate exemplars of “female engineering students” more strongly with exemplars of what “most people dislike”—especially in engineering departments in which females comprised less than 20% of the class, but also that the more negative these female students’ implicit normative evaluations of female engineers became, the more likely they were to be considering
dropping out of engineering (Spencer, Peach, Yoshida, & Zanna, 2010). These findings are exciting because we now have an implicit “thermometer” of the engineering culture—and, thus, can begin to study exactly what (presumably) subtle cues actually create a chilly climate. These findings are also exciting because they led to a logical, follow-up intervention study in which we created a “belongingness” intervention (patterned after Walton & Cohen, 2007) designed to convey the message that female students do, in fact, “belong” in engineering (i.e., we employed audio-taped interviews with several upper-year male and female students that conveyed the message that although most students worry during their first year about whether they belong in engineering, these worries lessen with time). This intervention was delivered at the beginning of the fall semester. So far, after 3 years of data collection, we have discovered that the belongingness intervention increases female engineering students’ grades, prevents their implicit norms of women in engineering from becoming negative, and increases their intentions to remain in engineering (Logel, Walton, Peach, Spencer, & Zanna, unpublished data)! Interestingly, it also prevented male students’ implicit norms of women in engineering from becoming negative over time. In future interventions, it might make sense to “beef up” the belongingness intervention to make the normative belief even more clear that both female and male engineering students not only discover that they personally belong in engineering, but also come to recognize that all admitted students, both male and female, belong in engineering.

Self-Esteem and Health

Another recent line of research in which I am currently interested concerns the negative relation between self-esteem and health. Put simply, individuals low in self-esteem experience poorer health. Danu Stinson, Christine Logel, and I thought this might be so because individuals with low self-esteem experience poorer close relationships (i.e., they incorrectly believe their relationship partners do not like/respect them) and that poorer close relationships are likely to relate to poorer health. So far, in two longitudinal studies, we have demonstrated that self-esteem assessed early in the term negatively predicts health outcomes later in the term, and that this effect is mediated by the perceived quality of relationships assessed at a time half-way between the assessments of self-esteem and health outcomes. In one of the studies we even found that change in self-esteem predicted change in health outcomes mediated by change in the perceived quality of close relationships (Stinson et al., 2008).

Having demonstrated that low self-esteem individuals experience poorer health because they erroneously perceive their relationships to be problematic, we next created an intervention, a “psychological flu shot” if you will, designed to improve their perceptions of their relationships and, thereby, their health outcomes—and, ultimately, their self-esteem. Specifically, we employed a
standard self-affirmation manipulation (in which participants select the value that is most important for themselves and explain why this value is so important). Preliminary results suggest that the self-affirmation manipulation entirely eliminated the negative relation between self-esteem and health by improving the health outcomes among low self-esteem participants who reported high levels of stress—and that this effect seemed to be mediated by decreasing the perception of stress in their relationships (Logel, Stinson, Shepherd, & Zanna, unpublished data).

**Scaled-Up Interventions in Appropriate Settings**

Are the sort of intervention studies described above sufficient to change the world? Obviously, they are not. This implies that, if we “bottle” a theoretically based intervention in either a laboratory or field experiment, we ought to consider the next step of disseminating the intervention more widely, perhaps with more than a little help from colleagues who have, at present, more expertise in disseminating such interventions. For example (and to be clear about what I am suggesting), once we discover (and replicate the findings) that female engineering students come to implicitly feel they do not belong in engineering and once we discover that a belongingness intervention works to prevent this from happening, we (with the help of others) should try to make this intervention an integral part of freshmen orientation week in our university, then in all universities in Canada, and then in all universities in North America, etc. For quality control, we should also continue to evaluate the effectiveness of the intervention as it is disseminated more widely. Of course, this is more than any individual faculty member can accomplish, but it is something for the field to contemplate and something to work toward in the future—in the service of making the world a better place.

**There is Nothing So Practical as a Good . . . Design**

I also believe that, as experts in research methodology, we can also improve the quality of assessing intervention experiments and, more generally, of policy-relevant research by creating better research designs. Put simply, designs that capture the applied question at hand can be incredibly useful and practical. Let me now turn to two examples: one ongoing and one from the distant past.

**Evaluating Tobacco Control Policies with Quasi-Experimental Designs**

Just as we now believe that medical, and even clinical, treatments ought to be evidence based, so should we expect health policy to be evidence based. In recent years, Geoff Fong and his many colleagues (including me) have been evaluating the national tobacco control policies (such as banning “light/mild” brand labels, creating more effective warning labels, banning cigarette advertising, increasing
cigarette taxes) mandated by the Framework Convention on Tobacco Control, the first international treaty devoted to health. Because governments, not researchers, control the implementation of such policies, we obviously cannot randomly assign one or more countries to “roll out” a particular policy. Thus, we have to use a quasi-experimental design that best approximates a true experiment. Following Campbell and Stanley (1963), we are tracking a panel of smokers in several countries, including Canada, United States, United Kingdom, and Australia. When one country (e.g., United Kingdom) initiates a national tobacco control policy (e.g., bans light/mild brand labels, or mandates new, more effective warning labels) between waves of data collection, it becomes the “experimental” country and the other countries (i.e., Canada, United States, and Australia) become the “control” countries with respect to that particular tobacco control policy. Thus, we have designed a pretest–posttest, longitudinal panel study with a comparison between the country that initiates a new, national tobacco-control policy and the countries that do not (Borland et al., 2008).

So far, this quasi-experimental design comes as close as possible to the sort of experimental designs that existed in the 1960s when Campbell and Stanley wrote their classic article on quasi-experimental designs. But, interestingly and importantly, we do not design experiments nowadays like we did in 1960s. What is new? Put simply, there is much more emphasis on testing mediation hypotheses. That is, nowadays we want to know the mechanisms/processes by which the independent variable influences downstream dependent variables. In the present context, for example, how can we be sure that, if smokers in the United Kingdom (compared to the smokers in the three control countries) are more likely to intend to quit smoking following the initiation of new, improved warning labels, this “effect” is actually due to the warning labels—and not something else?

Our solution to this problem was to create a quasi-experimental design (circa 2000) in which we attempted to assess mediation variables uniquely relevant to each of the potential tobacco-control policies under investigation. So, notwithstanding our critique of the standard way mediation hypotheses are tested in experimental research (Spencer, Zanna, & Fong, 2005), we believe that, if the downstream effects of a given policy are mediated by processes uniquely relevant to the policy and, importantly, not by processes uniquely relevant to other policies, we have strong, discriminative evidence that the initiated policy was, indeed, the variable responsible for producing the downstream effects (Fong, Hammond, & Zanna, 2006).

Evaluating Smoking Prevention Programs. . .While Waiting for Nature to Take Her Course

So far, I have discussed relatively recent research conducted within the last decade. Let me now close by going “back to the future” to talk about research I
conducted over 25 years ago—and, in the process, finally explain the title of this essay.

The story begins when I was invited to collaborate on an ongoing smoking prevention program with colleagues in the Health Studies Department at the University of Waterloo. My colleagues in Health Studies had initiated a second-generation smoking prevention program in several school districts in Ontario. In addition to “beefing up” the intervention itself, the main change from the first-generation smoking prevention programs was that the Waterloo intervention was delivered earlier (in the sixth grade) before kids were likely to have started smoking in the first place. The logic was simple: If you want to prevent smoking, intervene before kids start smoking.

OK, but why did these colleagues want me to collaborate with them, given the facts that their intervention was, if anything, likely to be an improvement, that their logic was entirely reasonable and, importantly, that they had just finished collecting a posttest wave of data at the end of the school year? It took me some time (perhaps, a half dozen research meetings over a 2-month period) to figure out not only that my colleagues’ logic did, in fact, result in a design problem that put their future funding at risk, but also that, although they needed some help to solve the problem, they were reticent to explicitly ask for advice. Put simply, if you deliver the intervention early—when few kids are already smoking in the control schools—you will not be able to detect an effect of your intervention in the experimental schools. I came to realize that I was being asked to help design a “quick fix” to this problem in the face of an imminent site visit from the funding source!

So, while waiting for nature to take her course (which in this context means, while waiting for kids to start smoking), what did we do to convince the funding source to continue funding the project because the intervention was, in fact, making a difference (or, at least, was likely to make a difference)? To answer this question I came up with two design solutions: A correlational (“at risk”) design and an experimental (social psychological) design. (Because we did not have time to conduct the experimental design before the site visit, let me first describe the at-risk design.)

The at-risk design recognizes the fact that although most kids have not started smoking by the end of the sixth grade, some may have. In this situation, the question becomes: Can we identify these kids a priori? Although this was not my area of research, I did not have to be a “rocket scientist” to suggest that the kids who were most at-risk to start smoking were probably those whose parents, siblings, and/or friends smoked. And, because my colleagues had collected this pretest data, we were able to create an at-risk index. Two relatively straightforward questions followed this insight: (1) Are the at-risk kids (even if only 10–15% of the sample) randomly assigned to the control schools, in fact, smoking by the end of the sixth grade?, and (2) If so, are the at-risk kids who were randomly assigned to the experimental schools smoking, too, or (hopefully) not? Well, it turns out
that the at-risk kids in the control schools had started to smoke and did so to a significantly greater extent than the at-risk kids in the experimental schools. So, it looked like the intervention was working—and, happily, was working for the kids for whom we most wanted it to work, i.e., those who were the most at-risk (see Flay et al., 1985, for a report of the completed study).

A couple of years later my colleagues conducted a follow-up smoking prevention intervention (in which they manipulated whether researchers, teachers, or school nurses delivered the intervention) that allowed us to carry out an experimental design in the tradition of social psychology. The cover story for the experiment was that, because summer was around the corner and summer camp directors were in the process of hiring summer camp counselors, we were interested in seeing whether camp directors would recommend hiring the same or different camp counselors than prospective campers, such as our participants. So, students were led to believe we were evaluating several applicants for the job of camp counselor. In the critical videotaped job interview the applicant either smoked a cigarette or not. Thus, in a $2 \times 2$ design, students from experimental versus control schools evaluated a smoking versus nonsmoking applicant for a summer camp counselor job. Interestingly—and importantly—before kids had started to smoke (and, thus, before we could detect whether the intervention had prevented kids from smoking) we found that sixth graders evaluated the smoking job applicant more negatively, but only if they had attended an intervention school (Towson, Pepperall, & Zanna, 1984).

Although nowadays we could (and probably would) assess students’ implicit attitudes toward smoking (as a way to rule out demand characteristics as an alternative explanation), I still believe it is important to consider the sort of experiment we conducted “while waiting for nature to take her course.” In fact, if I had to do it all over again, I’d make the dependent variable more consequential by convincing participants that their evaluations of the camp counselors would actually influence who would be hired!

In conclusion, in the spirit of Kurt Lewin, although I truly hope we never stray too far from doing curiosity-based, basic experimental research in the lab and in the field that is inspired by social issues, I also hope that we begin to change our academic culture to become a whole lot more serious about mobilizing our knowledge—by creating and testing theory-based interventions that we eventually disseminate—to make the world a better place—and, at least, to make our granting agencies happy, and, more importantly, to inspire our graduate students! And while we are at it, when we evaluate social policies, let us make sure we use “state of the art” designs.

References


MARK P. ZANNA is a University Professor and Chair of Psychology at the University of Waterloo. He received his bachelor’s and doctoral degrees from Yale University. His area of research is the psychology of attitudes. Currently, he is studying self-esteem and prejudice at both the explicit and implicit levels. A former President of the Society of Experimental Social Psychology (1985) and the Society of Personality and Social Psychology (1997), he currently edits Advances in Experimental Social Psychology (since 1991) and the Ontario Symposium on Personality and Social Psychology (since 1981). A Fellow of the Royal Society of Canada, he is the recipient of Distinguished Scientific Contribution Awards from the Canadian Psychological Association (D. O. Hebb Award in 1993), the Society of Personality and Social Psychology (D. T. Campbell Award in 1997), the Society of Experimental Social Psychology (in 2007), and the Society for the Psychological Study of Social Issues (K. Lewin Award in 2010).