

II Instructor's Manual

Table of Contents

1	<i>Introduction</i>	GM11
1.1	<i>Game synopsis</i>	GM12
	<i>Concurrent conversations</i>	GM13
2	<i>Model Schedules</i>	GM17
2.1	<i>Standard Schedule</i>	GM17
	<i>Schedule of Game Sessions</i>	GM17
2.2	<i>Expanded Schedule</i>	GM18
2.3	<i>Compressed Schedule</i>	GM18
2.4	<i>Long Class meetings</i>	GM18
3	<i>Roles and Factions</i>	GM19
3.1	<i>List of roles and factions</i>	GM19
	<i>Factions</i>	GM19
	<i>Secret societies</i>	GM19
	<i>Players</i>	GM20
4	<i>Game Setup</i>	GM23
4.1	<i>Overview</i>	GM23
4.2	<i>Rules and Procedures</i>	GM23
	<i>Committees</i>	GM24
	<i>Credibility Points</i>	GM28
	<i>Victory Objectives</i>	GM28

4.3 *Sample rubrics for grading* GM30

Writing GM30

Speaking GM32

5 *Game Management* GM33

5.1 *Narrative* GM33

Conference Schedule for 1971 GM36

Prior to 1971 GM38

Sample Conference Schedule for 1971 GM39

Sample Conference Schedule for 1972 GM41

Sample Conference Schedule for 1973 GM43

Sample Conference Schedule for 1974 GM45

Sample Conference Schedule for 1975 GM47

Proposals that may come up according to game play GM49

5.2 *Items you might need* GM50

Credibility Points GM51

APA positions costs GM52

Establishing Liminality GM53

Learning Objectives GM54

Conceptual issues at stake GM59

Readings GM63

Overview of events and assignments in the game GM68

Initial credibility of each character GM73

Committees, Publications and Exhibits GM74

Labs and possible research proposals GM76

5.3 *Research Committee Proposals* GM78

Emotional responses in a human child. GM79

Mental rotation of three-dimensional objects GM81

On Being Sane in Insane Places GM82

Adjustment of Homosexual and Heterosexual Women GM84

A Re-analysis of the Reliability of Psychiatric Diagnosis GM86

Interpersonal Dynamics in a Simulated Prison GM88

Small World GM91

5.4 *Game-Master Interventions* GM91

- Possible Mod: DSMIV compromise* GM91
- Poor Presidential tone* GM92
- Not grasping importance of DSMIII* GM92
- Failed Research cycle* GM92
- Innateness* GM92
- Possible Modification: Voting* GM93

6 *Debriefing* GM95

6.1 *What really happened* GM95

- Actual presidents (Psychological association):* GM95
- Actual history* GM95

6.2 *After this era* GM102

- Margaret Mead debunked* GM102
- Are Ebbinghaus' experiments replicatable?* GM102
- Abortion* GM102
- The DSM-V Controversy* GM103

7 *Works Cited* GM105

8 *Indicies* GM109

Instructors Manual Appendices GM113

- A *Real-world papers* GM113
 - Mental rotation instructors]Miller, George* GM113
 - Adjustment of Homosexual and Heterosexual Women* GM118
 - On being sane in insane places* GM122
 - Interpersonal Dynamics in a Simulated Prisoninstructors]Zimbardo, Philip: Research* GM137
 - A Re-analysis of the Reliability of Psychiatric Diagnosis instructors]Spitzer, Robert* GM167

B Lab Activities GM176

- Accuracy of Observation* GM176
Operant Conditioning: Role in Human Behavior GM179
Operant Conditioning Demonstration GM181
Defense Mechanisms GM184
To Sleep, Perchance to Dream GM186
Mental Illness GM190
Classical Conditioning GM193
Human Operant Conditioning GM196
Bringing the Clinic into the Undergraduate Classroom GM199
Discovering the Relationship Between Operational Definitions and Interobserver Reliability
GM207

Information Processing Capacity: A Visual Demonstration of the Magical Number Seven
GM215

- The Role of Prior Information in Dream Analysis* GM221

C Further Reading GM224

- DSM-III* GM224

D Obituaries and Tributes GM225

- Anne Anastasi* instructors]Anastasi, Anne GM225
George Albee instructors]Albee, George GM225
Irving Bieber instructors]Bieber, Irving GM225
Kenneth Clark instructors]Clark, Kenneth GM225
John Fryer instructors]Fryer, John GM226
Barbara Gittings instructors]Gittings, Barbara GM226
Harry Harlow instructors]Harlow, Harry GM226
Harold Lief GM226
Charles Socarides instructors]Socarides, Charles GM226
John P. Spiegel instructors]Spiegel, John GM226

List of Tables

2.1 Outline of Game sessions	GM18
3.1 Initial Factions	GM19
3.2 Secret and Informal Factions	GM20
3.3 Character assignments for small class	GM21
3.4 Character assignments for large class	GM22
4.1 Proposals for Mental Illness	GM29
4.2 Sample rubric for research reports	GM30
4.3 Sample rubric for writing papers	GM31
4.4 Sample rubric for speaking	GM32
5.1 Schedule for first conference	GM37
5.2 Major events of 1971	GM39
5.3 Possible events of 1972	GM42
5.4 Major events of 1973	GM43
5.5 Major events of 1974	GM45
5.6 Major events of 1974	GM47
5.7 Leona Tyler Principle (proposed by L. Tyler, when president)	GM49
5.8 Position Statement: Hospital Privileges for Psychologists (opposed by G. Albee)	GM49
5.9 Position Statement: Affirmative Action (Proposed by K. Clark)	GM50
5.10 Credibility Points	GM51
5.11 Credibility costs for service to the APA	GM52
5.12 Excerpt from APA Guidelines for an Undergraduate Major in Psychology, 2007	GM54
5.13 Excerpt from APA Guidelines for an Undergraduate Major in Psychology 2.0, 2013	GM55
5.14 Timeline of critical events	GM57
5.15 Major Issues for debate	GM62
5.16 Core text reading list	GM64
5.17 Overview of reading assignments, by character	GM68
5.18 Overview of specific assignments, by character	GM71
5.19 Overview by topic	GM72

5.20 Elections to be held each year GM73

5.21 Initial credibility GM74

5.22 Likely Actions by Committee GM75

5.23 Possible lab activities GM77

6.1 Real APA (psychological) Presidents GM95

List of Figures

- 5.1 Advertisement for Librium, 1962. From GM53
- 5.2 Advertisement for Valium, 1965. GM53
- 5.3 Advertisement for Blue Cross, 1960. GM54

- 6.1 Manuscripts and Archives Division, The New York Public Library.
(1972 - 1972). *Gittings, Kameny, Dr. John E. Fryer, Dr. Judd Marmor on panel* Retrieved from [urlhttp://digitalcollections.nypl.org/items/510d47e3-830b-a3d9-e040-e00a18064a99](http://digitalcollections.nypl.org/items/510d47e3-830b-a3d9-e040-e00a18064a99) GM97
- 6.2 Screenshot of message posted on NARTH homepage after death of
Nicolosi GM98

1 Introduction

Welcome to “Defining the Mind.” This game is set around the creation of the DSM-III in the 1970’s. Prior to this period, psychoanalysis dominated psychiatry and behaviorism dominated psychology, both to the point that change seemed unthinkable for many.

This is not a game about the accumulation and use of power in the ordinary way. This is a game about academic credibility and evidence. In the ideal of academic life towards which we strive, these are one and the same. In reality, of course, they are not.

But there are moments—Kuhn used the political metaphor of a revolution—where known anomalies in the dominant paradigm become too significant to ignore, and shifts occur. In these moments, academic credibility takes on a new form. And like political revolutions, new movements in science gain credibility by appeal to external entities and ideologies. Indeed, most of the debates between factions in these periods regard the Philosophy of Science, not the evidence *per se*.

The writing of the DSM-III marks one of these moments of revolution in Psychiatry.

The DSM-I’s “Revised Nomenclature” is 39 pages long (Included as Appendix APA *Diagnostic and Statistical Manual Mental Disorders*, 1952 section G on p. 452). The DSMII’s is 40 pages long (Included as Appendix Socarides, S. “Homosexuality and Medicine”, 1968 section I on p. 619). The equivalent section (2) of the DSMIII is 300 pages long. The DSMIII-R added another 250 pages or so and the DSMIV another 100. Since then, the DSM IV-TR and the DSM V have continued to grow. There is as massive difference in scale in these versions, but it is the difference *in kind* between the taxonomy used in early DSMs and III that interests us in this game. And while that taxonomy has grown since the III, it has not changed in nature.

WE DO NOT EXPECT undergraduate students to produce a diagnostic taxonomy in this game. Rather, we are expecting students to determine upon what basis such a diagnostic taxonomy should be based: observable symptoms or hypothetical causes. These are questions of

Philosophy, not necessarily questions of Psychology or Psychiatry. And we throw in Thomas Szasz and his sympathizers to drive home the importance of having a definition in the first place.

1.1 *Game synopsis*

The narrative of the game begins, but *does not end with*, the demedicalization of homosexuality. This game is not intended to be a political game wherein the players are centrally interested in the collection and use of power. Rather, this game is meant to be about evidence, and the interpretation of evidence.

Bieber claims, for example, to have about 30% success rate in his 'treatment' of homosexuality. Is this an adequate 'success' rate given the struggles the other 70% of his patients report—as experienced first-hand by Ron Gold? Or even compared against the societal stigma that is placed on homosexuals because of the 'procedure'?

More importantly, this game is not centrally *about* homosexuality. That issue is completed largely in the first session. This game is about whether or not science can legitimately posit underlying mechanisms to explain behavior or is restricted to merely cataloging and describing behavior and behavior correlations. Spitzer's proposed revisions to the DSM largely banish psychoanalytic taxonomies (and their assumed underlying psychodynamic mechanisms) from the official language of psychiatry. At the same time, Chomsky and Miller are reintroducing hypothetical mechanisms to Psychology, which has resisted them for half a century of behaviorism.

The three main factions have distinctly different views on this question:

- The psychoanalysts believe that one cannot categorize human behavior without understanding the underlying mechanisms.
- The behaviorists believe that science requires that they limit themselves only to describing behavior.
- The cognitivists offer a third way, a kind of hybrid between the two, where mechanisms are required for adequate explanation of behavior, but they are ontologically limited to those that can be realized in the physical substrata of the human brain.

After the initial storm over homosexuality, the central question of the game is scientific legitimacy of psychoanalysis. This is a complicated issue, and will make allies of those initially divided over homosexuality—Marmor and Speigel will need to work with Socardes /Bieber to protect psychoanalysis from Spitzer and his allies.

Spitzer himself is a bit of a wild card. In reality, Spitzer was never enthralled with psychoanalysis. Once he understood the negative effects that medicalization had on homosexuals, he consolidated control and forced the change to the DSM through sheer force.

To this day, advocates for NARTH and 'reparative therapy' complain (probably justly) about his willingness to bend the rules in pursuit of his ends. But this doesn't mean that he was wrong—the evidence produced by Socades and Bieber is terrible, while Hooker's evidence is quite strong. It is up to you and the student to decide if your Spitzer will rely on the power of good ideas and evidence or ram the changes through using parliamentary tricks.

There are a number of other issues that will arise, including whether there can be, or should be, psychological study of women and minority communities. And almost as importantly, what the responsibilities to the public good a scientist has.

The game culminates in the vote defining 'mental illness.' There should be at least 3 candidate definitions proposed: by Spitzer, Szasz and the Psychoanalysts. In reality, no one definition ever passed, and the Psychiatric and Psychological community still operate without a formal definition. In the world of your game, it's pretty much up in the air if anyone will succeed in getting their definition passed.

Concurrent conversations

There are three main 'tracks' or 'conversations' happening concurrently in this game.

1. The first—the dominate narrative—starts with the demedicalization of homosexuality (1 below), follows the Spitzer proposal to rewrite the DSM-III (3 below) and ends with the definition of mental illness (5 below).
2. The second regards research proposals, labs and ethics (2 below). Proposals are reviewed weekly by the Research Committee, and if approved, the experiment performed by the members of the class in the next class session.
3. The third regards the policies of the APA, specifically the role in the wider culture of the APA and Psychology and Psychiatry broadly construed (4 below). The elections to the board as well as Presidential elections are included in this conversation, even as they impact the other two. This thread climaxes in the debate on the Leona Tyler principle and Chomsky and Clark's arguments against.

These are three major topics /events that every conference should address. The first, demedicalization, is really a sample question to set up the major issues of the game.

1. Demedicalization of Homosexuality. (week 1)

The demedicalization of homosexuality is almost totally assured. The only characters who argue against it are Bieber and Socarides. Their reliance on psychoanalytic techniques will not sway Marmor or Hopcke, who are committed to demedicalization. This may be frustrating to your students. It is wise to remind them that the main question of the game—the legitimacy of psychoanalysis as a scientific enterprise—is very much in play, and if they strategize correctly, they can win the game even though they lost the first battle.

The NY Times obituary of Socarides quotes Gilbert Herdt of National Sexuality Resource Center in San Francisco as saying “Socarides outlived his time.” That is roughly correct. Many psychiatrists in 1971 may not have noticed that the medicalization of homosexuality was a problem for the homosexual community, but once it was pointed out, the opinion swung dramatically: 58% of the population approved of the demedicalization in 1973. And today there is great embarrassment about that era.

As such, the main debate here is about the procedure by which homosexuality will be removed. Green and Marmor will be tempted to just ram it through the board of directors without consulting the population. That should be avoided. I've set up the game so that Marmor's original proposal should be remanded to the nomenclature committee, which will report in 1972, requiring a final vote of the membership in 1973. While the result is not in doubt, without that process, Socarides' rhetoric that the decision was political, not scientific, will be bolstered.

1. Research. (throughout)

The ‘normal business’ of the APA should continue throughout. Research proposals should be brought forward, the best one approved and the study conducted (that is, as much as is practical). There are two types of proposals that can be brought to the Research committee: real, historical experiments that I have ‘anonymized’ for use in the game, and in class-room experiments from the APA's (*Psychological Activities Handbook for the Teaching of Psychology*). I've selected experiments that reflect the major issues in the game. They are available in Section Labs and possible research proposals subsection 5.2 on p. GM76.

1. Association Business /DSM-III. (week 2–4)

The second section of the game addresses the central question of the legitimacy of psychoanalysis. It does so tangentially by asking

students to debate the evidence that should be used to form the taxonomy of mental illness as contained in the DSMIII. Students should not be debating whether this or that criteria is included, but rather whether the DSMIII ought to be structured around the psychodynamic hypothesis or the study of observable symptoms. You could say the debate is whether mental illness is classified by its causes or its effects.

At the same time, the Spitzer taskforce should propose a covering definition of 'mental illness' that will directly challenge the views of Szasz and Albee. It should fail (as it did in reality).

1. The social responsibilities of social scientists. (Week 3–5)

Leona Tyler is prompted to propose the 'Leona Tyler Principle' to the board in week 3. That should spark Ken Clark and Noam Chomsky to object, and propose a symposium in the following year. Of course, these can come at any time, but they are vital to the end game.

1. End Game: Definition of 'mental illness' and Fission of the APA. (week 5)

Members of the factions have the option to form splinter associations. In reality:

Behaviorists found the "Society for the Experimental Analysis of Behavior" (original founding 1957), and founded two journals: *Journal of the Experimental Analysis of Behavior* and *Journal of Applied Behavioral Analysis*.

Psychoanalysts form the American Psychoanalytic Association (actual founding in 1911, but grew rapidly in this era) and the journal *Journal of the American Psychoanalytic Association (JAPA)*, and

Cognitivists form the Cognitive Science Society (1979) and its associated journals.

No students are given direct instruction on whether the fission requires withdrawal from the APA. Some, esp. those playing Beiber and Socarides, may feel stronger about the fission than the cognitivists Miller and Chomsky. This is historically appropriate—by the mid-1980s, cognitivism generally dominated the APA. And contemporary behaviorism has adapted to resemble cognitivism in many ways, by including talk of 'motivations' and of the 'function of behavior' in their explanations.

2 Model Schedules

2.1 Standard Schedule

I've taught 'Defining the Mind' as a Tuesday-Thursday 1.5 hour per class section. That gave me the structure to require proposals be distributed on Tuesday for discussion on Thursday.

The schedule I have used is represented in table Outline of Game sessions Table 2.1 on p. GM18. Replicated here from Section 3.4 on page 101.

Schedule of Game Sessions

Year—Location	
Session	Activities
1971—Washington DC	
A	Presidential Address George Miller Symposium: "Psychiatry: Friend or Foe to Homosexuals: A Dialogue," (Dr H. Anonymous, E. Hooker) T. Szasz "The Myth of Mental Illness" Presentation of 'mental rotation' task: gamemaster
B	Marmor "Limitations of Free Association" Proposal from J. Marmor Proposal from C. Socarides . Petition from G. Albee. Research report from G. Miller on 'mental rotation' task
1972—Dallas	
A	Presidential Address Albert Bandura Symposium on Medical Model (G. Albee, T. Szasz) Report from taskforces
B	R. Spitzer 'The Fiegner Criteria' P. Gebhard on the Kinsey reports H. Harlow 'Lust, latency and love' Research Report

1973—Honolulu	
A	Presidential Address Paper(s) Reports from taskforces
B	Symposium Proposal to create "Spitzer Taskforce" [other proposals] Research Report
1974—Philadelphia	
A	Presidential Address Symposium Paper(s)
B	Open hearings on proposed definition of 'mental illness' [other proposals] Research Report
1975—Chicago	
A	Presidential Address Open vote of the membership on definition of mental illness. Paper(s)
B	Symposium: [other proposals] Research Report

Table 2.1: Outline of Game sessions

2.2 *Expanded Schedule*

2.3 *Compressed Schedule*

2.4 *Long Class meetings*

3 Roles and Factions

3.1 List of roles and factions

Factions in this game are relative ‘loose.’ The basic theoretical frameworks in place have factional alignment, but there are plenty of independents who ‘lean’ towards one or more factions.¹

More importantly, there is a large set of topics upon which the factions are neutral. Some factional members, such as Marmor and Bieber will find themselves as opponents on major issues.

Factions

There are three factions, representing the three major theoretical frameworks in psychology in the 1970s.

¹ We cover—very briefly—James’ functionalism and Titchener’s structuralism, though not quite using those names, in the Game book, section Brief history of the concept of ‘Psychology’ section 2.2 on p. 23. These views are not given voice in the game, as contemporary parallels tend to reside in Philosophy departments, not psychology.

Psychoanalysts	Cognitivists	Behaviorists	Independents
I.Bieber /C. Socarides, MD	G. Miller, PhD	A. Bandura, PhD	T. Szasz, MD
J.Marmor, MD	N. Chomsky, PhD	A. Bandura, PhD	G. Albee, PhD
J. Spiegel, MD	D. Marr, ABD	H. Harlow, PhD	J. Fryer, MD
H. Lief, MD		E. Hooker, MD	K. Clark, PhD
R. Green, MD		A. Anastasi, PhD	
R. Hopcke, MD		L. Tyler, PhD	
		R. Spitzer, MD	
		P. Gebhard, PhD (Anthro)	
		F Kameny /B Gittings (Activists)*	
		D. Fordney-Settlage, MD*	
		K. Freund, MD*	
		J. Piaget, PhD*	
		R. Gold (Journalist)*	
		S. Milgram, PhD*	
		P. Zimbardo, PhD*	

Secret societies

The ‘secret societies’ ought to meet outside of class, to maintain their anonymity. They have no objectives other than their own, but they

should orchestrate their actions so they help each other out where necessary.

For example, Ron Gold is charged with 'helping' Spitzer draft the proposal to demedicalize homosexuality. I encourage the whole GayPA to have a version drafted *prior* to needing it in 1973.

GayPA (secret)	Young Turks (secret)	Sexual Research (informal)
J. Fryer, MD	R. Hopcke, MD*	H. Leif, MD*
E. Hooker, PhD	J. Spiegel, MD	R. Green, MD*
R. Gold*	J. Marmor, MD	D. Fordney-Settlage, MD*

Table 3.2: Secret and Informal Factions

Players

Between 16 and 26. Every character has specific assignments in writing, politics and research. Table Character assignments for small class Table 3.3 on p. GM21 summarizes these assignments.

Well-known publications by each character are included in Section Brief Sketch of the Game Characters section 4.2 on p. 140 of the gamebook.

Cast of 16

Row	Name	Area	Game Play
1	Robert Spitzer	Independent - Psychiatrist	Nomenclature committee chair in 1971, petitions for task force in 1973.
2	George Miller	Cognitivist	President in 1971, leader of cognitivists.
3	Anthony Bandura	Behaviorist	Vice-president in 1971, behaviorist, but may join cognitivists.
4	Harry Harlow	Behaviorist	Former Pres. of APA, leader of behaviorists, defends aversion therapy '74.
5	Noam Chomsky	Cognitivist	Critic of behaviorism, founder of cognitivism, social activist. Symposium '74, debate with Piaget '75.
6	Leona Tyler	Independent - psychologist: counseling	(run for) Vice President in 1972, install the 'Leona Tyler Principle' as president '73.
7	Anne Anastasi	Independent - psychologist: psychometrics	Run for VP in 1973, form the committee on women in Psych, and the committee on the Psych of Women. Neutralize the gender-biased language of the official APA calls for papers, '71.

8	John P. Spiegel	Psychoanalyst	Run for VP in 1975, reliable partner of Spitzer, straight advocate for the GayPA. Propose ('71) and complete a report ('72) on homosexuality in psychology and psychiatry.
9	Evelyn Hooker	Behaviorist	Early studies of homosexuality (1953), the intellectual 'grandmother' of the current movement. Calming influence on the GayPA, and scientifically reliable source for their arguments.
10	George Albee	Independent - clinical psychologist	Medical model symposium '72, propose clinical psych health care system.
11	Ken Clark	Independent - psychologist	Run for VP in 1974, expert testimony in Brown v. Board of education. Symposium '74.
12	Judd Marmor	Psychoanalyst - Freudian	Propose removal of homosexuality '71, condemn Socarides JAMA paper '72.
13	Thomas Szasz	Independent - psychiatrist	Critic of Psychoanalysis, and more broadly, the medicalization of psychiatry. Paper '71.
14	Irving Beiber / Charles Socarides	<i>can be split</i> Psychoanalyst – Freudian	Classical psychoanalyst, specializing in 'treatment' of homosexuality. God-father of 'reparative therapy' movement, as he is the mentor of contemporary Nicolosi (NARTH).
15	John Fryer	Independent - psychiatrist	Dr. H. Anonymous.
16	Kurt Freund	Independent - psychiatrist	Propose rejection of aversion therapy '74.

Table 3.3: Character assignments for small class

Cast of 20

To add additional characters to the game:

- add Paul Gebhard /Harold Lief, Independent - Anthropologist / Psychoanalyst - Jungian&Anthropologist representative on Spitzer Task Force /Jungian psychiatrist.
- add Richard Green, MD, expert in transgenderism, student of John Money, who was also on the actual Spitzer task force (Money may be added to future versions of the game).
- then Ron Gold, Journalist, Activist. Convinces Spitzer that classification is doing more harm than good. Should also participate in a Symposium in 1973 giving a talk titled "Stop it, you're making me sick."

- Split Socarides/Bieber into two roles.
- Split Gebhard/Lief into two roles.

Cast of 27

And then add to this:

Row	Name	Faction / Views	Game Play
17	Paul Gebhard / H. Lief	<i>can be split</i> Independent - Anthropologist / Psychoanalyst - Jungian	Anthropologist representative on Spitzer Task Force / Jungian psychiatrist.
18	Robert Hopcke	Psychiatrist - Jungian	Jungian psychiatrist who updates theory to respect homosexuals. Historically inaccurate.
19	Jean Piaget	Psychologist – Developmental	Old man at this time. Debates Chomsky on innateness in 1975.
20	Dr. Fordney Settlage	Gynecologist	Member of the Spitzer Task Force, critic of androcentrism of psych. / psychiat.
21	Kameny / Gittings	'Homophile' Activists	Activists, co-founders of Mattachine Society of Washington DC.
22	Marr	Cognitive Scientist	Young researchers, articulates the 'levels' of explanation of cognitive science.
23	Milgram	Psychologist	Presents 'obedience' study, proposes 'small world' study.
24	Zimbardo	Psychologist	Proposes and defends the prison study.

Table 3.4: Character assignments for large class

4 Game Setup

4.1 *Overview*

While there is a great deal to prepare for this game, my experience has taught me that the largest challenge the students will face is understanding that psychology and psychiatry were not always as they are now presented.

Most students come to the game believing that (a) psychology is a science and (b) psychiatry is a medical practice. Neither of those claims were settled in the public mind in 1970. In fact, most of the tension in this game revolves around the efforts to make psychiatry conform to ‘the medical model’ and psychology conform to the model of the natural sciences, thereby legitimizing them as worthwhile endeavors.

Psychology, on the other hand, had been suffering a crisis of legitimacy since its inception, which is why so many of their arguments are really metatheoretical arguments on the nature of science.¹ The behaviorists believe that introducing new methodologies, like the cognitivists propose, would further weaken their claims to be a scientific discipline.

¹ As noted in the introduction to the Instructor’s manual.

4.2 *Rules and Procedures*

Generally speaking, our students come to psychology and psychiatry through textbooks. And as a result, they are primed to believe, unreflectively, that contemporary narrative of these disciplines is settled fact. Furthermore, these students know nothing of the rise and regularization of health insurance companies and billing procedures following the inception of Medicare and Medicaid in 1965. For them, managed care, billable hours and check-box diagnoses have always been a part of their medical experience. It was not in 1970.

Since the 1970s, the psychiatrists and clinical psychologists have been under extraordinary pressure to create a system of diagnosis that will allow them to be compensated for their work under the system of managed care. Spitzer’s shift from etiological psychiatry to descriptive

psychiatry was a major step in that struggle.

It is difficult to get them to feel the pressure to legitimize psychiatry and psychology that drove much of the work in this era. You can help set the stage by hanging posters of advertising for pharmaceuticals and health insurance policies from the era. I've highlighted a few examples in Section Ads for Psychopharmaceuticals and Health Insurance subsubsection 5.2 on p. GM53 of the Instructor's manual.

Many of the primary sources from this period—especially those from the members of the Spitzer task force—recall a time of great confusion. Spitzer is often characterized as acting almost single-handedly, ignoring the hard work that others put into the classifications they proposed. If students feel confused and overwhelmed with all of the proposals being brought forward and the changes being made, that is partially intentional.

While the removal of homosexuality from the DSM was the correct decision, both morally and scientifically, that does not mean that the dissenters who point out the political bullying that went into passing the resolutions do not have a point. Everything was up for grabs, and legitimacy had to be established by extra-scientific methods. Spitzer (the real Spitzer) used the power of Robert's rules and the unlimited mandate of his task force.

We want the students to use things more sophisticated, like evidence and reason. But it is probably enough if they come away from the game with the sense that this period in the history of psychology and psychiatry was an all out scramble for legitimacy. Having a sense of confusion and chaos may be a part of that experience.

Committees

Getting the committees to function well can be challenging—in the game as in real life. In the ideal form, the committees would do their work outside of class, and report in class to the Board at the end of each conference. In reality, you may have to give time at the start of class to allow committees to meet and hash out their decisions.

Board of Directors

Responsibilities:

- Hold open meetings each conference where topics can be discussed and voted upon.

Powers:

- Issue public proclamations on behalf of the membership.
- Full control of official publications, such as the DSM.

- Create ad-hoc committees and task forces, as necessary.
- Oversee and receive reports from the standing committees.
- Censure—can strip any member of credibility at any time. Generally reserved for use of ad hominem attacks or other bad behavior. Any number of credibility points can be stripped.
- Banning—rarely used, but available if necessary. Can place a life-time ban on any member at anytime.
- Confer 1–10 credibility points on retiring board members in recognition of their ‘distinguished service’

Potential Pitfalls

- The election process of VP’s may confuse students. It is not an uncommon practice in higher ed, but in my experience, the students find it very strange to elect VPs, not Presidents. It is probably worthwhile explaining it in person before game play begins.
- The game is setup to require a ‘built in’ 48-hour rule, where proposals to the board cannot be voted on in the same class period where they were presented. If that does not work for your class schedule, make sure everyone on the board and the membership understands it before the first open meeting.
- The boards I have worked with are tentative about using their powers, and are naturally conservative in taking actions. This, again, isn’t that surprising. For example, I’ve never had a board that was ready to act on Marmor’s proposal in 1971. They *always* want to come to consensus, and turf it to a committee that comes back in 1973. If the board decides to act immediately, that’s fine—you can flip the 1973 symposium on its head and have Socarides and Bieber present a panel on why they believe the demedicalization was incorrect, and have Gold as a critic.

Research Committee

The research committee is charged with distributing grants to fund research as well as enforcing the APA’s Ethical Standards in the practice of Psychology and Psychiatry (see Appendix F). **The ‘research’ thread of the game depends entirely on the research committee functioning well.**

Responsibilities

- Solicit proposals in the form of a ‘call for research grants’ — an example is included in Schedule of Game Sessions subsubsection 3.6 on p. 109 in the Game Book.

Powers

- Award grants to those proposals it deems excellent. Grants give the grantee the right to ‘use’ the class for their experiment for no more than 1 class session.
- Hearing and deciding on cases of research ethics, including recommending punishments to the Board for violators up to and including banning from the APA for life.
- Confer 1–5 credibility points on retiring board members in recognition of their ‘distinguished service’

Potential Pitfalls

- No one submits research proposals: we have included a number of sample proposals that are summaries of famous experiments from this era in Section Research Committee Proposals section 5.3 on p. GM78. We've also included a handful of classroom ‘experiments’ from the APA Handbooks on teaching that are consistent with the game and time period in the appendix. Table Possible lab activities Table 5.23 on p. GM77 lists labs and who might submit them as proposals for what purpose. Feel free to submit one or two of these to the research committee ‘anonymously’, each week, and if approved, assign one of the committee members to lead the study.
- Research committee fails to recognize ethical violations: This is especially important if Zimbardo submits his proposal. According to the 1968 guidelines, the proposal should be approved—as it was at Stanford. But there are obvious problems. If they approve it, the game master can intervene as “an outraged public” and demand a correction to the 1968 guidelines.

Nomenclature Committee

The Nomenclature committee is charged with maintaining the official terminology of psychology and psychiatry. This is embodied by the Diagnostic and Statistical Manual, which is the definitive source for definitions and classifications of mental disorders.

Responsibilities

- Maintain the official diagnostic and statistical manual of mental illness.

Powers

- Define what kind of behaviors qualify as ‘mental illnesses’, thereby (because of the rise of health insurance and managed care) defining

what kind of behaviors psychologists and psychiatrists can get paid to treat.

- Add a paper, panel or symposium topic to any conference agenda without review of the Program Committee.
- Confer 1–5 credibility points on retiring board members in recognition of their ‘distinguished service.’

Potential Pitfalls

- Spitzer will push the board to dissolve the Nomenclature committee and replace with a task force. If unsuccessful, all attempts to define ‘mental illness’ will be brought to the Nomenclature committee instead of to the Board. That will require that the Nomenclature committee exercise its power to add a session at an upcoming conference. The Game master will be required to ensure there is adequate time for discussion and voting on these contentious issues. If Spitzer is successful in his proposal to dissolve the Nomenclature committee, that debate will occur during the Board meeting where the Task Force reports.

Program Committee

The Program Committee is charged with creating the schedule for each conference, including disseminating the Call for Proposals and judging submissions.

Responsibilities

- Solicit proposals from the membership.
- Create the schedule for each conference.
- Confer 1–3 credibility points on retiring board members in recognition of their ‘distinguished service.’

Powers

- Decide who gets to speak at any conference, thereby determining who has the ability to gain credibility.

Potential Pitfalls

- **Disengaged committee:** in my experience, this committee has been the most difficult. It is hard for them to see that their power is probably the most significant in the game, as they control who is eligible to earn credibility, and hence, who is eligible to hold any of the positions in the APA. Students are often inclined, as they probably should be, to simply allow anyone who wants to talk to

do so. You may wish to restrict the number of presentations per session, in order to help them understand the importance of the committee's decision-making.

- **Overly strict committee:** I've never had this happen, but it is worth noting it as a possibility. Every student is required to present at least once in the course of the game. A Program committee could, in theory, 'blacklist' a player, thereby making it impossible for that student to complete his or her goals. The game master needs to balance the principles of academic respect with the politics of the game. If this were to happen, the best solution would be for the blacklisted student to use the game mechanisms to fix his or her problem: run for the committee, or recruit an ally to do so, and allow his or her paper at the next conference. If that is impossible—say the blacklisted student waited until the final week to recognize the problem—you'll need to use your own discretion.

Credibility Points

At the beginning of each conference (each 'week' on the standard schedule), distribute 1 credibility point to every student entering the game. Each student *must* give this point to the individual that he or she believes gave the best speech, research report or symposium presentation by the end of the week.

The character of Ron Gold, as a journalist, starts the game with 10 credibility points that he *must* distribute at the beginning of the game.

Running for office 'costs' credibility. Table Credibility Points Table 3.2 on p. 93 in the Game Book shows the suggested costs of running for a seat. You are welcome to adjust this table for your own purposes, but if you do so, make sure to provide a copy to all the students. A blank version is included on page GM52 of the Instructor's Manual.

Credibility can be given or traded for favors, **but**, like a block-chain, all credibility transfers must be recorded in public by the board of the APA. If a senior member of the APA, say A. Bandura, wants to lend some of his credibility to an issue pushed by a junior member with less credibility, he is welcome to do so, **but** he must formally declare the transfer of credibility during a Board meeting of the APA so that everyone is aware of said transfer.

Victory Objectives

The final vote—and consequently 'victory' in the game—is regarding the definition of 'mental illness.' Table Proposals for Mental Illness Table 4.1 on p. GM29 of the Game Book lists the possible definitions

and likely supporters.

Proposal	Those 'normally' affiliated
APA Task Force	Spitzer, Gebhard, Leif, Green
Psychoanalytic	Bieber / Socarides, Marmor, Speigel, Hopcke
Behaviorist	Bandura , Harlow, Hooker
Szasz	Szasz, maybe Albee

Table 4.1: Proposals for Mental Illness

APA Task Force (in Spitzer's role sheet)

A Medical disorder is a relatively distinct condition resulting from an organismic dysfunction which in its fully developed or extreme form is directly and intrinsically associated with distress, disability, or certain other types of disadvantage. The disadvantage may be of a physical, perceptual, sexual, or interpersonal nature. Implicitly there is a call for action on the part of the person who has the condition, the medical or its allied professions, and society.

A mental disorder is a medical disorder whose manifestations are primarily signs or symptoms of a psychological (behavioral) nature, or if physical, can be understood only using psychological concepts. (1978, p. 18)

Behaviorist

A person can be called 'mentally ill' when he or she exhibits emotional or behavioral functioning which is so impaired as to interfere substantially with his or her capacity to function in society.

Psychoanalytic

A person is mentally ill when he or she suffers from internal conflicts that may be subconscious or unconscious, manifesting behavior that is unwanted or disturbing to the individual or the society.

Szasz (and maybe) Albee

There is no 'thing' called 'mental illness,' only sets of behaviors that may be destructive to an individual and his or her society.

The final vote is tallied in credibility, not 'votes'. To win, a definition *must* have more than 50% of the credibility available. If no definition surpasses this value, the APA can move forward without a definition, or it can change the rules for voting and vote again—for example, hold a run-off between the top-two vote getters.

The victory objectives can also be "modded" to a compromise. See Section Possible Mod: DSMIV compromise subsection 5.4 on p. GM91 of the Instructor's Manual.

4.3 Sample rubrics for grading

Writing

For research reports:

Standards	Exemplary	Satisfactory	Unacceptable
Presentation of Evidence. (See item 3 in Learning Objectives of the Game item 13 on p. GM59)			
The student identifies any possible confounding variables that are controlled in the experiment, and can explain *how* they are controlled. The student identifies the relationship identified between the independent and dependent variable, and is able to explain the statistical test proficiently.	The student correctly identifies the hypothesis being tested, the experimental design and the independent variable and dependent variable.	The student misidentifies the hypothesis, the variables and or the experimental design.	
Interpretation of results. (See item 3 in Learning Objectives of the Game item 13 on p. GM59)			
The student is able to relate this result to other results presented in class—is this finding consistent with other data presented so far?	The student correctly interprets the results: was the hypothesis confirmed? Disconfirmed? Etc.	The student is unable to connect the results of the study to the hypothesis being tested, or any overarching theoretical positions being explored.	
Application to theoretical positions in game play. (See item 6 in Learning Objectives of the Game item 13 on p. GM59)			
The student is clearly connects the experiment directly to the major theoretical frameworks in game play, demonstrating understanding of more than one position.	The student makes a connection, but it isn't clear or it isn't direct.	The student makes no connection	

Table 4.2: Sample rubric for research reports

For writing Tasks

For other writing tasks, the objectives are a bit more abstract:

Standards		
Exemplary	Satisfactory	Unacceptable
Legitimacy		
Student raises, or preemptively addresses (depending on role), questions regarding the legitimacy of evidence. (See item 2 in Learning Objectives of the Game item 13 on p. GM59)		
Context		
Where appropriate, student is able to contextualize views under discussion into their historical context (See items 1 and 4 in Learning Objectives of the Game item 13 on p. GM59).	Student recognizes context, but does not use it in analyses.	Student treats all theories as authoritative.
Primary sources		
Student uses historically appropriate primary sources (See item 5 in Learning Objectives of the Game item 13 on p. GM59)	Student cites historically appropriate primary sources, but does not engage with them	Student does not include historically appropriate primary sources – sources not primary, or not historically appropriately.
Implications		
Student is able to distinguish between the major theoretical positions represented in the game, and interpret evidence as bearing on those positions in some way. (Item 6 in Learning Objectives of the Game item 13 on p. GM59)	Student can 'rehearse' distinctions between theoretical positions, but cannot connect theories with evidence	Student mistakes or confuses theoretical positions.

Table 4.3: Sample rubric for writing papers

Based on Clabough and Clabough 2016

Speaking

Attached is a rubric borrowed from the National honors society for Human sciences. I prefer it for a science-based speaking assignment to others I have seen, as it emphasizes clarity of content over rhetoric.

Standards		
Exemplary	Satisfactory	Unacceptable
Organization		
Has a clear opening statement that catches audience's interest, maintains focus throughout, summarizes main points	Has opening statement relevant to topic and gives outline of speech, is mostly organized, provides adequate "road map" for the listener	Has no opening statement or has an irrelevant statement, gives listener no focus or outline of the presentation
Content		
Demonstrates substance and depth, is comprehensive, shows mastery of material	Covers topic, uses appropriate sources, is objective	Does not give adequate coverage of topic, lacks sources
Quality of conclusion		
Delivers a conclusion that is well documented and persuasive	Summarizes presentation's main points, draws conclusions based upon these points	Has missing or poor conclusion, is not tied to analysis, does not summarize points that support the conclusion
Delivery		
Has natural delivery, modulates voice, is articulate, projects enthusiasm, interest, and confidence, uses body language effectively	Has appropriate pace, has no distracting mannerisms, is easily understood,	Is often hard to understand, has voice that is too soft or too loud, has a pace that is too quick or too slow, demonstrates one or more distracting mannerisms
Response to Questions		
Demonstrates full knowledge of topic, explains and elaborates on all questions	Shows ease in answering questions but does not elaborate	Demonstrates little grasp of information, has undeveloped or unclear answers to questions

Table 4.4: Sample rubric for speaking

Adapted from <https://rubrics.kon.org/about.html> (National Honors society for the Human sciences)

5 Game Management

Each week represents one annual conference. At each conference, there are a number of things that need to happen, although the order in which they happen is up to you. These are:

- President addresses the whole.
- Board meeting considers any proposals, votes.
- Committees meet, and report to the Board if necessary.
- Research papers presented.

For each week, you will need to prep and bring with you:

- Credibility points: print Table Credibility Points Table 5.10 on p. GM51 of the Instructor's Manual
- Proposals for the research committee (Table Possible lab activities Table 5.23 on p. GM77), if there are no proposals from students.
- The primary source of any historical studies (Table Historical studies subsubsection 5.2 on p. GM77) of experiment that will be reported.

Specific items needed for each conference are listed on the individual conference pages: pages GM39-GM47.

When and if the Research Committee approves labs that can actually be carried out, you will need to work with the student to ensure that required materials are available.

5.1 Narrative

This game is about the crisis of legitimacy faced by psychology and psychiatry in the 1970's. This crisis came on two fronts: the external front and the internal front. The external crisis included the gray rights protests and ultimate demedicalization of homosexuality, the rise of psychopharmaceuticals, and the Goldwater affair and the institution of the Leona Tyler principle. Internal crisis includes Marmor's

open worries about the legitimacy of free association, Rosenhan's famous 'Being Sane in Insane Places' study, Thomas Szasz, and the rise of cognitivism and the death of behaviorism.

In game play, these two crises are separated into the two basic 'modalities' of play: politics and research. These are orchestrated to be mutually reinforcing, where lab experiences inform discussion of policy, and questions about policy drive research. That all depends, of course, on timing. And timing in a reacting game is unpredictable.

I've tried to suggest years in which different characters ought to offer political proposals and research grants so the narrative arc is coherent. But it is easy to fall out of step. No single research experience is 'mission-critical' to finishing the political narrative, but the student experience will generally be enhanced if the schedule is adhered to.

THE POLITICAL: The dominant narrative of the game starts with Marmor's shocking and radical proposal to drop Homosexuality (302.0) from the DSM-II in 1971 (Marmor Proposal 1 on p. R136). This is met with a counter-proposal from Socarides /Bieber to study the whole taxonomy of 302 ('Sexual Deviations') and report back with a proposal of their own (Socarides Proposal 1 on p. R123). *That* is met with a counter proposal from Spiegel and Green to study the history and effectiveness of 'treatment' of homosexuals (Spiegel Proposal 1 on p. R140 and Green Proposal 1 on p. R108).

As game master, you know that Spiegel and Marmor are collaborating as members of the Young Turks, but Socarides /Bieber and the members of the Board in 1972 do not.

Spitzer's movement in 1973 to both remove Homosexuality and rewrite the DSMIII is one of the points in this game where the political and the research intersect. Motivating Spitzer is the notion that taxonomy ought to be based on evidence, not etiology. Rosenhan, included in game play in 1974 (Fryer Research 1 on p. R41, provides some of that evidence. But his tactics are political—a decision that would open rhetorical strategies to his opponents well into the 2010's.

Other proposals in this thread include Anastasi's proposals to start a division on the psychology of women(Anastasi Proposal 2 on p. R27), Clark's proposal to enact an affirmative action policy (Clark Proposal 1 on p. R32), the Leona Tyler Principle (Tyler Proposal 1 on p. R80) and the concluding symposium with on the social responsibilities of social scientists.

THE RESEARCH: The secondary narrative (B-Plot) of the game revolves around the research committee. Research proposals are made, and if successful, the student gets access to the class as subjects. The researcher is bound to report on their proposal in the following con-

ference.

Many of the research tasks specified are classics in the field, or easy in-classroom activities taken from the APA handbooks. Major Issues for debate Table 5.15 on p. GM62 shows how research tasks relate to the political throughout the class.

The game master ought to distribute the *real* paper relating to each lab activity **after** it has been presented by the student. These are listed in each year's schedule below, and included on the GitHub site for the book.

The twist to the research committee is Zimbardo. If Zimbardo proposes the prison study in 1973 or 1974, and if it is approved by the research committee, the game master ought to intervene and claim (as the high voice of reason) that the experiment is clearly unethical. As the research committee was probably correct in assessing that there was nothing in Zimbardo's proposal that violated the 1968 ethical standards (APA "Ethical Standards of Psychologists" 1968 section F on p. 446), the game master ought to charge the research committee with drafting a proposed revision to the APA Ethical Standards.

THE FINAL VOTE: The final vote on the definition of mental illness these threads together. In one view, this is merely a nomenclature matter for a scientific community—a physicist would not care if their definition of 'force' matched the public views or not. On another view, this definition, as it is put forward by an authoritative body, will profoundly impact the lives of millions of people: whether or not they can get insurance coverage for treatment, whether or not their views can be dismissed as those of a 'diseased' mind, etc.

Voting the this final case is done by casting credibility points. As such, the result of the game ought to be wholly determined by game play. The psychoanalysts obviously have a advantage in sheer numbers of players, but I believe they will have a deficit of credibility if things follow the pattern of play I've experienced in my classes.

Conference Schedule for 1971

Presidential Address:

Dr. G. Miller "The Future of Psychology"

Distribution of proposals to be considered this year:

- J. Marmor: proposal to remove 'homosexuality' from the DSMII (302.0).
- C. Socarides & I. Beiber: proposal to create taskforce on sexual deviation.
- J. Spiegel and/or R. Green: proposal to create task force of historical study and literature review of homosexuality in psychology and psychiatry.

Symposia

"Psychiatry: Friend or Foe to Homosexuals: A Dialogue."

- Dr. E. Hooker "The mental health of non-patient male homosexuals."
- Dr. H. Anonymous, "I am a homosexual and a psychiatrist."
- F. Kameny and/or B. Gittings "Gay, Proud and Healthy."

Research Papers

- Dr. T. Szasz "The Myth of Mental Illness."
- Dr. J Marmor "Limitations of Free Association."

Open Board Meeting

General business meeting agenda:

Committee Reports

- Dr. Tyler (Research)
- Dr. Spitzer (Nomenclature)
- Dr. Hooker (Conference)

Old Business

New Business

Discussion and vote on:

- Proposal from J. Marmor
- Proposal from C. Socardies /I. Beiber
- Proposal from G. Albee

Nominations and elections for:

- Vice President 1972
- Replacement for Milgram, member at large on the Board of Directors

Table 5.1: Schedule for first conference

Prior to 1971

It's important that the first week set the pattern for all weeks after. Each conference should have research reports and research proposals. *Proposals* propose research to be carried out by the class. *Reports* report to the class on the result of that research. So if a student *proposes* a research task in 1971, he or she will *report* on that research in 1972.

That means we need to set up a research task prior to the meeting of the class in 1971, so it can be *reported*. I've given the task of reporting on 'Mental Rotation of 3-dimension objects' to George Miller. I've chosen this task because it is a classic of cognitive psychology, and there is a freely-available demonstration on the APA's website.

The **best** approach here would be to register for an account here: <https://www.apa.org/ed/precollege/ptn/2018/09/online-laboratory-renovation>, and have all the students in the class participate in the experiment, then collect the data and pass it to the student playing Miller.

If it is impossible to do this, because of class schedule or internet access or whatever, have the 'Miller' student read the actual paper—which is included in the 'Primary Source' directory of the game archive. The student ought to present the paper data *as if* it was class data.

Sample Conference Schedule for 1971

Item	First session	Second session
Presidential Address	Miller "Psychology as a means of promoting human welfare" (See Miller Writing Assignment 1 on p. R177 of the Role Sheet book)	
Symposium	"Psychiatry: Friend or Foe to Homosexuals: A Dialogue," Fryer (Fryer Writing Assignment 1 on p. R41 of the Role Sheet book), Hooker (Hooker Writing Assignment 1 on p. R161 of the Role Sheet book), Kameny (Kameny Writing Assignment 1 on p. R63) / Gittings (Gittings Writing Assignment 2 on p. R66 of the Role Sheet book)	
Proposals for the Board	Removal of homosexuality (Marmor Proposal 1 on p. R136 of the Role Sheet book)	<ul style="list-style-type: none"> • Proposal from C. Socarides to create task force on sexual deviation. (Socarides Proposal 1 on p. R123 of the Role Sheet book) • Proposal from J. Spiegel and R. Green to create task force to conduct a study of the history of the treatment of homosexuality in psychology and psychiatry. (Spiegel Proposal 1 on p. R140 and Green Proposal 1 on p. R108 of the Role Sheet book)
Proposals for Research Committee		<ul style="list-style-type: none"> • "Adjustment of Homosexual and Heterosexual Women" (Gittings Research 1 on p. R66) (proposed by Barbara Gittings, if in game, if not, gamemaster). Sample is included as Adjustment of Homosexual and Heterosexual Women subsection A on p. GM118 of the Instructor's Manual. • "Emotional responses in a human child" (proposed by gamemaster). Sample is included as Emotional responses in a human child subsection 5.3 on p. GM79 of the Instructor's Manual. The original paper is available on psych classics: http://psychclassics.yorku.ca/Watson/emotion.htm
Research Reports		<ul style="list-style-type: none"> • Miller (Miller Research 1 on p. R177), presenting Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. <i>Science</i>, 171, 701-703. Paper included in the Appendix as Mental rotation subsection A on p. GM113 of the Instructor's Manual.

Table 5.2: Major events of 1971

Notes:

Of course, the Little Albert experiment cannot be completed by actual students. If the research committee approves the experiments, the instructor should charge a student with presenting the original papers during the following years' conference.

Students may recognize the "little albert" experiment, although I don't use the name 'Little Albert' in the proposal. They will also probably find the experiment unethical. Many people do. The version of the proposed experiment I've included herein contains no plan for deconditioning Albert. See the Little Albert subsubsection 6.1 on p. GM95 for a discussion of the history of this famous experiment.

Papers to be distributed for next week:

- 1971—Following approval of Miller Research 1 on p. R177 of the Role Sheet book:
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703. Included as Mental rotation of three-dimensional objects subsection 5.3 on p. GM81 in the Instructor's Manual.

Sample Conference Schedule for 1972

Papers /Proposals other than the Presidential address are *likely* to be presented, but all depend on game play.

Item	First session	Second session
Presidential Address	Bandura , "Behavioral theory and models of man" (Bandura Writing Assignment 1 on p. R148 of the Role Sheet book)	
Symposium	On the 'medical model' (Szasz Writing Assignment 1 on p. R99, Marmor Writing Assignment 1 on p. R137 and Albee Writing Assignment 1 on p. R19 of the Role Sheet book)	
Proposals for the Board	Proposal to recognize the contributions of Evelyn Hooker (Gebhard Proposal 1 on p. R50)	<ul style="list-style-type: none"> • Proposal from Nomenclature Committee to write a new DSM according to description of symptoms, not causes (Pro: Spitzer Proposal 2 on p. R92 and Opposed: Lief Proposal 1 on p. R116) • Proposal from Hopcke to limit medical care of mentally distressed individuals in hospitals to licensed medical doctors (i.e. psychiatrists) (Hopcke Proposal 1 on p. R112 of the Role Sheet book) • Counter proposal from Albee to limit the role of psychiatrists to the admission of drugs. (Albee Proposal 1 on p. R19 in the Role Sheet book) • Fordney-Settlage may be able to broker a compromise. (Fordney-Settlage Proposal 1 on p. R36 in the Role Sheet Book)
Proposals for Research Committee		<ul style="list-style-type: none"> • Proposal from Fordney-Settlage to study the sexual experience of teenage girls seeking contraception for the first time (Fordney-Settlage Research 1 on p. R36 in the Role Sheet book).

Research Reports	<p>Depending on outcome of 1971 proposals, one (or all) of:</p> <ul style="list-style-type: none"> • Gittings Writing Assignment 1 on p. R66, "Adjustment of Homosexual and Heterosexual Women" • Distribute the actual report of the 'Little Albert' experiment. Paper is available on psych classics: http://psychclassics.yorku.ca/Watson/emotion.htm • Socarides on Sexual Deviation (Socarides Writing Assignment 1 on p. R124) • Green (Green Writing Assignment 1 on p. R108) and Spiegel (Spiegel Writing Assignment 1 on p. R140) on history of homosexuality, with responses from Bieber (Bieber Writing Assignment 1 on p. R131), Marmor (Marmor Proposal 2 on p. R136)
------------------	---

Table 5.3: Possible events of 1972

Other papers to expect

- Paper by Spitzer on Feighner Criteria (Spitzer Writing Assignment 1 on p. R92 of the Role Sheet book), Anastasi in support (Anastasi Writing Assignment 1 on p. R25 of the Role Sheet book)
- Paper by Gebhard on the Kinsey report (Gebhard Writing Assignment 1 on p. R50 of the Role Sheet book)
- Paper by Harlow "Lust, latency and love" (Harlow Writing Assignment 1 on p. R152 of the Role Sheet book)

Other possible proposals**Research committee Issues****Papers to be distributed (if activities are approved)**

- 1972—Following research report by Gittings Research 1 on p. R66 of the Role Sheet book:
 - Siegelman, "Adjustment of Homosexual and Heterosexual Women" *The British Journal of Psychiatry* (1972) 120: 477–481. Included in appendix as Adjustment of Homosexual and Heterosexual Women subsection A on p. GM118 in the Instructor's Manual.

Sample Conference Schedule for 1973

Papers /Proposals other than the Presidential address are *likely* to be presented, but all depend on game play.

Item	First session	Second session
Presidential Address	Depends on previous elections: Tyler , Clark possible	
Symposium	Nomenclature committee-organized open panel discussion of the proposed removal of 'Homosexuality' from 302.00. Should include Gold Writing Assignment 1 on p. R56 Kameny/Gittings , Green, Spitzer and Socarides/Bieber to be panelists, but they will speak extemporaneously, not deliver formal papers.	
Proposals for the Board	<ul style="list-style-type: none"> • Removal of homosexuality immediately (Spitzer Proposal 1 on p. R91 of the Role Sheet book) • Proposal from Green create a journal titled <i>Archives of Sexual Behavior</i> (Green Proposal 2 on p. R108 in the Role Sheet book) • Spitzer task force, dissolution of nomenclature committee (Spitzer Proposal 1 on p. R91 of the Role Sheet book.) 	(likely / possible): <ul style="list-style-type: none"> • Creation of the committee on Women in Psychology (Anastasi Proposal 1 on p. R26 in the Role Sheet book) • Creation of the Society for the Psychology of Women (Anastasi Proposal 2 on p. R27 in the Role Sheet book)
Proposals for Research Committee	<ul style="list-style-type: none"> • Prison study. Included in appendix as Interpersonal Dynamics in a Simulated Prison subsection 5.3 on p. GM88 in the Instructor's Manual. (proposed by Zimbardo, if in game, if not, gamemaster) (Zimbardo Research 1 on p. R103 in the Role Sheet book) • "A re-analysis of the reliability of psychiatric diagnosis," Included in appendix as A Re-analysis of the Reliability of Psychiatric Diagnosis subsection 5.3 on p. GM86 of the Instructor's Manual. (proposed by Spitzer Research 1 on p. R92) . 	
Research Reports	<p>Depending on outcome of 1972 proposals, one (or both) of:</p> <ul style="list-style-type: none"> • Fordney-Settlage "Sexual experience of younger teenage girls seeking contraceptive assistance for the first time." (Fordney-Settlage Research 1 on p. R36 in the Role Sheet book) 	

Table 5.4: Major events of 1973

Other papers to expect

- Paper by Tyler on individual differences and their importance for counseling—likely Presidential Address. (Tyler Writing Assignment 1 on p. R80 of the Role Sheet book).

Other potential proposals

- Marmor - if not in 1972: official condemnation / censure of Socarides JAMA paper as nonscientific and a "Monstorous attack on homosexuality." (Marmor Proposal 2 on p. R136 of the Role Sheet book)
- Leona Tyler principle (Tyler Proposal 1 on p. R80 of the Role Sheet book).

Papers to be distributed (if activities are approved)

- On Gay Pride day in 1973, someone firebombed the 'UpStairs Lounge', a gay club in New Orleans, killing 32 people, 3 of whom were never identified. The Game Master should distribute the news prior to the APA conference. See <http://www.patheos.com/blogs/friendlyatheist/2013/06/24/remembering-the-upstairs-lounge-the-u-s-a-s-largest-lgbt-massacre-happened-40-years-ago> for a blog-length history, or "Let the Faggots Burn: The UpStairs Lounge Fire" by Johnnny Townsend for a book-length history.
- 1973—Following research presentation of Szasz Research 1 on p. R99 of the Role Sheet book:
 - Rosenhan 'On being sane in insane places' Science v.. 179 (Jan. 1973), 250–258. Included in appendix as On being sane in insane places subsection A on p. GM122 in the Instructor's Manual.

Sample Conference Schedule for 1974

Papers /Proposals other than the Presidential address are *likely* to be presented, but all depend on game play.

Item	First session	Second session
Presidential Address	Depends on previous elections, but likely: Clark, Anastasi	
Symposium	TBD	
Proposals for the Board	TBD	TBD
Proposals for Research Committee	<ul style="list-style-type: none"> • Proposal from Milgram (Milgram Research 1 on p. R70 in the Role Sheet book). If Milgram is not a character, proposal should come from gamemaster See Small World subsection 5.3 on p. GM91 for sample proposal. Original paper included in supplementary materials on GitHub.) • Proposal from Fryer to replicate Rosenhan in the classroom (Fryer Research 1 on p. R41). 	
Research Reports	Depending on outcome of 1973 proposals, one (or both) of:	
	<ul style="list-style-type: none"> • Zimbardo Research 1 on p. R103 • Spitzer Research 1 on p. R92 	

Table 5.5: Major events of 1974

Other papers to expect

- Symposium on “Social Responsibility of Intellectuals” — Proposed by Clark Clark - 1974 Writing Assignment 2 on p. R32
 - Clark - 1974 Writing Assignment 1 on p. R32
 - Chomsky Writing Assignment 2 on p. R169
- Obedience - Milgram Writing Assignment 1 on p. R71
- Genetic hypothesis and diagnostic taxa - Leif Writing Assignment 1 on p. R116

Open hearings /symposium on the Spitzer Task Force definition of ‘mental illness’.

Other possible proposals

- Deem aversion therapy to treat homosexuality immoral. Freud Proposal 1 on p. R47

In opposition: Harlow.

Papers to be distributed (if activities are approved)

- Following research presentation by Zimbardo Research 1 on p. R103 of the Role Sheet book:
 - Haney, C. Banks, C. & Zimbardo, P. (1973). "Interpersonal Dynamics in a Simulated Prison," International Journal of Criminology and Penology 1, p. 69–97, Included in appendix as Interpersonal Dynamics in a Simulated Prison subsection A on p. GM137 in the Instructor's Manual.
- Following research presentation by Spitzer Research 1 on p. R92 of the Role Sheet book
 - Spitzer, R. and Fleiss, J. (1974). "A Re-analysis of the Reliability of Psychiatric Diagnosis" Brit. J. Psychiat. 125, 341–7. Included in appendix as A Re-analysis of the Reliability of Psychiatric Diagnosis subsection 5.3 on p. GM86 in the Instructor's Manual.

Sample Conference Schedule for 1975

Papers /Proposals other than the Presidential address are *likely* to be presented, but all depend on game play.

Item	First session	Second session
Presidential Address	Depends on previous elections, but possible: Spiegel, Clark, Anastasi	
Symposium	TBD	
Proposals for the Board	TBD	TBD
Proposals for Research Committee	TBD	
Research Reports	Depending on outcome of 1974 proposals, Milgram Research 1 on p. R70 and /or Fryer Research 1 on p. R41	

Table 5.6: Major events of 1974

Other papers to expect

- Open forum on definition of mental illness (called by Nomenclature /Spitzer Taskforce) (See Spitzer Proposal 4 on p. R92).
- Zimbardo on the 1968 code of ethics - Zimbardo Writing Assignment 1 on p. R104.
- Fordney-Settlage on the disparity between psychoanalytic theories of male and female sexuality - Fordney-Settlage Writing Assignment 1 on p. R37.
- Hopcke on Jungianism and homosexuality - Hopcke Writing Assignment 1 on p. R112.
- Marr on levels of explanation in Psychology - Marr Writing Assignment 1 on p. R172.
- Symposium on 'innateness' between Piaget (Piaget Writing Assignment 1 on p. R75) and Chomsky (Chomsky Writing Assignment 1 on p. R169) with implications for social responsibility and homosexuality.

Other possible proposals

- Clark should propose affirmative action (See Proposals that may come up according to game play Table 5.9 on p. GM50 and Clark Proposal 1 on p. R32). Stanley Milgram should oppose (See Stanley Milgram, PhD chapter 11 on p. R69)¹

- Zimbardo and the Research committee - revisions to the 1968 code of ethics.

Papers to be distributed (if activities are approved)

- Following research presentation by Milgram Research 1 on p. R70:
 - Experimental Study of the Small World Problem. (1969). An Experimental Study of the Small World Problem, 32(4), 425–443.
<http://doi.org/10.2307/2786545?refreqid=search-gateway:a88fe1dd7db6336d575e302coc15221e>
Included in appendix as Small World subsection 5.3 on p. GM91
in the Instructor's Manual.

Proposals that may come up according to game play

As citizens, members of the APA have the right to advocate for any cause through the myriad of political advocacy organizations, but when psychologists and psychiatrists speak for the profession through APA public stances and proclamations, it should be from science and professional experience.

On occasion psychiatrists are asked for an opinion about an individual who is in the light of public attention or who has disclosed information about himself/herself through public media. In such circumstances, a psychiatrist may share with the public his or her expertise about psychiatric issues in general. However, it is unethical for a psychiatrist to offer a professional opinion unless he or she has conducted an examination and has been granted proper authorization for such a statement.

Table 5.7: Leona Tyler Principle (proposed by L. Tyler, when president)

POSITION STATEMENT (1974)

Approved by the Board of Trustees, December 1970

This statement was prepared by the Committee on Psychiatry and Psychology.

Because of professional and legal considerations, the ultimate medical responsibility for patients admitted to hospitals should remain with licensed physicians. Psychologists, like other nonmedical professionals, should be eligible for some type of hospital appointment.

Table 5.8: Position Statement: Hospital Privileges for Psychologists (opposed by G. Albee)

POSITION STATEMENT

Approved by the Board of Trustees, December 1977

Approved by the Assembly of District Branches, October 1977

This statement was prepared by the Committee of Black Psychiatrists and recommended by the Council on National Affairs

THERE IS a continuous need to increase the number of minority psychiatrists; the American Psychiatric Association has consistently demonstrated its commitment to the principle of affirmative action as reflected in its efforts of recruitment and training of minority psychiatrists. APA has previously developed and instituted policies recognizing and supporting the special mental health issues of minority populations; however, there are serious threats to affirmative action programs that have facilitated the following endeavors: APA reaffirms these commitments and policies by 1) issuing a public statement drawing attention to the potential deleterious effects that such threats pose to the delivery of health services to minority groups; 2) actively participating with other professional and educational groups to assure continued recruitment and training of minority candidates in medical disciplines; and 3) further exploring and developing, through its appropriate components, mechanisms to assure continued implementation of these commitments.

Table 5.9: Position Statement: Affirmative Action (Proposed by K. Clark)

5.2 *Items you might need*

Prior to 1971, you should provide the students with the mental rotation task, as discussed previously in Prior to 1971 subsection 5.1 on p. GM38.

You should probably have hard copies of:

- Sample conference schedule (Section Schedule of Game Sessions subsubsection 3.7 on p. 114 of the instructor's manual).
- Sample Call for Papers and Symposia (Section Schedule of Game Sessions Table 3.6 on p. 102 of the game book).
- Sample call for research grants (Section Schedule of Game Sessions subsubsection 3.6 on p. 109 of the game book).
- Any Primary source papers to be distributed (see table Possible lab activities Table 5.23 on p. GM77 of the instructor's manual)

Credibility Points

'Thumbs Up' image licensed CCo Creative Commons license, from
<https://pxhere.com/en/photo/941135>

 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point	Table 5.10: Credibility Points  You have my Approval 1 Credibility point
 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point
 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point
 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point	 You have my Approval 1 Credibility point

APA positions costs

Position	Cost
Vice-President of the APA	10
Member of the Board	5
Chair, Research Committee	5
Chair, Nomenclature Committee	5
Chair, Program Committee	3

Table 5.11: Credibility costs for service to the APA

Establishing Liminality

Ads for Psychopharmaceuticals and Health Insurance

To set the stage, you might want to print out some of these ads and hang them around the room. The observant student might note that most of these drugs are marketed for use on women, but the marketing is to the men—all the while being tested almost exclusively on men.

Many such can be found here: <https://prescriptiondrugs.procon.org/view.resource.php?resourceID=005597> (1960s) and here: <https://prescriptiondrugs.procon.org/view.resource.php?resourceID=005598> (1970s)

It might also be worth putting up some ads for health insurance, as this is important to understanding the drive to standardize treatment. These are not too difficult to find on line, but here's a link to a few: <http://www.vintageadbrowser.com/money-ads-1970s/8>

There are a great number of blogs and websites dedicated to storing and distributing these historical images. Here are a few:

- Miltown: <http://www.homeeverafter.com/miltown-a-piece-of-1950s-homemaker-history/>
- Tofranil: <https://www.biopsychiatry.com/imipramine/> 350x525xtofranil.jpg.pagespeed.ic.xDYlzJklSs.webp
 - <https://www.biopsychiatry.com/imipramine/tofranil.html>
- Librium: <https://www.mmm-online.com/channel/med-ad-hall-of-fame-to-induct-lerner-girgenti-and-rubin/article/155796/>
- Valium: http://faculty.weber.edu/ewalker/Medicinal_Chemistry/topics/Psycho/psycho.htm

At the same time, we should remember that Medicare and Medicaid were created in 1965, and the health insurance industry was undergoing a simultaneous revolution.

- BlueCross 1960: <http://www.decodog.com/inven/MD/md28530.jpg>



Figure 5.1: Advertisement for Librium, 1962. From <https://www.mmm-online.com/channel/med-ad-hall-of-fame-to-induct-lerner-girgenti-and-rubin/article/155796/>

Figure 5.2: Advertisement for Valium, 1965.

Learning Objectives

In many ways, this is not a game about Psychology or Psychiatry, it is a game about the Philosophy of Science. This shouldn't really be surprising, as the history of Psychology and Psychiatry is intimately linked to the history of Philosophy of Science. Arguments made by Wundt, Freud, Watson, Hull Skinner and Miller all rely heavily on claims regarding what is or is not a legitimate scientific claim.² This issue has not disappeared from undergraduate Psychological classroom either. The APA's guidelines for the Undergraduate Major in Psychology, published 2007, lists 10 learning outcomes for a major. The first is in box Learning Objectives Table 5.12 on p. GM54, and the second in Learning Objectives Table 5.13 on p. GM55.

Goal 1: Knowledge Base of Psychology

Demonstrate familiarity with the major concepts, theoretical perspectives, empirical findings, and historical trends in psychology

Suggested Learning Outcomes

- 1.1. Characterize the nature of psychology as a discipline
 - a) Explain why psychology is a science.
 - b) Identify and explain the primary objectives of psychology: describing, understanding, predicting, and controlling behavior and mental processes.
 - c) Compare and contrast the assumptions and methods of psychology with those of other disciplines
 - d) Describe the contributions of psychology perspectives to interdisciplinary collaboration.

Table 5.12: Excerpt from APA Guidelines for an Undergraduate Major in Psychology, 2007

If the reader compares to the definitions offered by the various historical figures in the 'Brief history of the concept of 'Psychology' section 2.2 on p. 23' in the gamebook, you will no doubt recognize the theoretical pluralism embodied in the APA's statements. These guidelines were superseded by the more prescriptive 'Version 2.0' in 2013, but the same themes reappear:

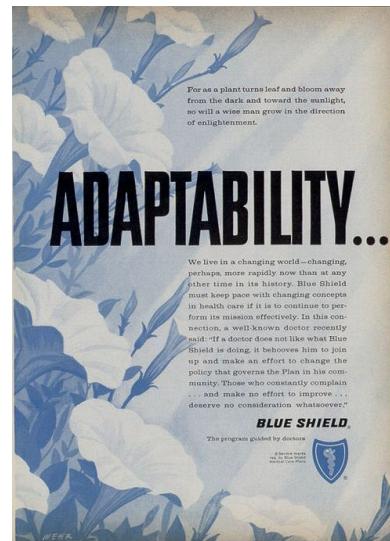


Figure 5.3: Advertisement for Blue Cross, 1960.

² See, e.g. Wundt (1902, pg 5–6) Freud XX, Watson (Mathematical paper), Skinner XX, Hull, (1935), Miller (citing Suppes)

Goal 1: Knowledge Base in Psychology

1.1 Describe key concepts, principles, and overarching themes in psychology.

1.1a Use basic psychological terminology, concepts and theories in psychology to explain behavior and mental processes.

1.1b Explain why psychology is a science with the primary objectives of describing, understanding, predicting and controlling behavior and mental processes.

1.1c Interpret behavior and mental processes at an appropriate level of complexity.

1.1d Recognize the power of the context in shaping conclusions about individual behavior.

1.1e Identify fields other than psychology that address behavioral concerns.

Table 5.13: Excerpt from APA Guidelines for an Undergraduate Major in Psychology 2.0, 2013

Version 2.0 instrumentalizes the ‘nature of psychology as a discipline’ into the measurable phrase “use basic psychological terminology, concepts and theories.” It is also notable that the phrase “and mental processes” appears in three of the subpoints (1.1a, 1.1b and 1.1c), rather than just one (b). One might be tempted to assert that behaviorism is well and truly dead in Psychology.

My main reason for quoting this document is, however, 2007 (a) and 2013 (1.1b).

The question “is psychology a science” is a question of demarcation – the classic issue in the Philosophy of Science. But answering *why* psychology is a science assumes an affirmative answer to the demarcation problem and, hence, a particular view in the Philosophy of Science. This is not a settled issue.

For example, Wundt charged Hebart with non-scientific investigations into introspection,³ because Hebart did not adequately control the environment. Watson charged that McDougall’s notion of behaviors having a ‘purpose’ was non-scientific,⁴ because McDougall did not follow strict verificationism in the Philosophy of Science. When Hull charged that Tolman’s explanations were unscientific because they posited internal entities, he did so by citing Newton⁵ (1935). Miller and Chomsky rejected Skinner’s arguments by citing Philosophers of Science such as Patrick Suppes, who contend that the history of science shows that scientists do posit internal entities, if they can be modeled mathematically.

³ Wundt, 1902, pg. 5–10

⁴ Watson 1929, pg. 25–6 “The Behaviorist finds no scientific evidence for the existence of any vitalistic principle, such, for example, as Prof. MacDougall’s ‘purpose.... .There are many things we cannot explain in behavior just as there are many things we cannot explain in physics and chemistry, but where objectively verifiable experimentation ends, hypothesis, and later theory, begin.’”

⁵ Newton, the critic would note, posited ‘force’, an invisible entity that was criticized at the time as an ‘occult’ power.

Textbooks

Most *Introduction to Psychology* textbooks address Goal 1 in Chapter 1. They all include a brief historical overview, and all note Wundt, James, Freud and Pavlov.

Book	Def. Of 'Science'	Historical highlights (I.e. Section heads)
Lilienfeld et al. 2014	"A systematic <i>approach</i> to evidence" (p.8)	Early History (27-29), Structuralism (30), Functionalism (30), Behaviorism (30-31), Cognitivism(31) and Psychoanalysis (31-32) Significant section on pseudoscience (p. 12-26)
Cacioppo and Freberg 2018	"a special way of learning about reality through systematic observation and experimentation." (p. 36)	Greek philosophers, British empiricists, ancient physicians, 17th and 18th century natural scientists, Helmholtz (all 6-10), Wundt (11-12), Gestalt (12), James, (12-14), Freud (14) & Humanistic Psychology (15), Behaviorists and Cognitive Revolution (16-21) Defines 'psychology' as 'science of mind' rather than 'science of behavior and mental processes' (p. 5)
Cervone and Caldwell 2015	'scientific methods refers to a broad array of procedures through which scientists obtain information about the world' (p. 9)—Collect evidence, record observations systematically, record how observations were made.	Aristotle & Buddha (20-22) Locke & Kant (22-24), Wundt (24-25), James (25-26), Structuralism & Functionalism (27), Psychoanalysis, Behaviorism (28), Humanistic (28-29), Information Processing and the Cognitive Revolution (29-30)
Gazzaniga 2018	'psychology: the study through research of mind, brain and behavior'—"gain accurate knowledge about behavior and mental processes only by observing the world and measuring aspects of it" (p. 30)	Descartes (p. 10-11), Wundt (11-12), Structuralism (12), Functionalism (12), Evolution, Adaptation and Behavior (13), Psychoanalytic Approach (14), Behaviorism (15), Gestalt (15-16), Humanistic (16), Cognitivism (16) Additional definition of 'scientific method' that combines method with goals.

Hockenbury, Nolan, and Hockenbury 2014	The scientific method is a set of assumptions, attitudes , and procedures that guides all scientists, including psychologists, in conducting research (16) – empirical evidence: "Verifiable evidence that is based upon objective observation, measurement and/or experimentation." (16)	Influence of Philosophy and Physiology (3-4), Wundt (4), Titchener (4-5), James (5-6), Freud (7-8), Watson (8), Rogers (9)
Myers and DeWall 2015	"less a set of finding than a way of asking and answering questions" (5)	Wundt (2), Structuralism and Functionalism (2-3), First Women in Psychology (3-4), Behaviorism (4), Freudian Psychology (5), Humanistic Psychology (5), Cognitive Revolution (5)
Zimbardo, Johnson, and McCann 2012	"based on objective, verifiable evidence—not just the opinions of experts and authorities, as we often find in non-scientific fields" (4)	Separation of Mind and Body and the Modern Biological Perspective (Descartes (12-13)), The Founding of Scientific Psychology and the Modern Cognitive Perspective (Wundt (13-14), Structuralism (Titchener) (14), James (14-15), Cognitive (15-16)), The Behavioral Perspective (16-17), The Whole-Person Perspective (17-19), The Developmental Perspective, The Sociocultural Perspective "Perspective" approach is unique in the set of textbooks reviewed.
Rathus 2012	"Psychology, like other sciences, seeks to describe, explain, predict, and control the events it studies... ...behavior and mental processes." (4)	Aristotle(8-9), Structuralism - Wundt (9), Functionalism - James (9-10), Behaviorism (10), Gestalt (10-11), Psychoanalysis (11-12)

Table 5.14: Timeline of critical events

When one reviews these offerings, one is struck by the lack of variation. All cite the same historical movements—including, somewhat surprisingly, humanistic psychology.⁶ They *all* define science as a method or procedure, and all reject the notion that science is a body of knowledge to be memorized. This cannot be random. No equivalent set of texts in Philosophy would be as uniform.

Moreover, the idea that science can be distinguished from non-science or pseudoscience based on method alone is highly question-

⁶ For example, Hothersall's *History of Psychology* covers all the views and thinkers covered here, right up to cognitivism, but contains no mention of Maslow or humanistic psychology. (Hothersall 1984)

able. Ptolemaic astronomy based its work on careful observation, and applied the rigorous methods of mathematics in order to describe and predict the movements of the heavens. So does astrology. On the other side of the dialectic, there is little in relativistic physics or quantum mechanics that is observable or was subject to experiment *before* it was accepted as scientific.

So neither way is the view articulated by these textbooks correct—there are non-scientific endeavors that fit the methodology-based demarcation, and there are scientific endeavors that do not. Simply speaking, one cannot demarcate science from non-science using methodology alone.

We are in dire need of a richer, more complete approach to introducing the scientific history of Psychology.

Learning Objectives of the Game

At the completion of this game, the students will be able to:

1. Recognize that the historical contingency of the concepts, principles and overarching themes that define and distinguish our contemporary disciplines that study the mind.
2. Identify and respond to historical examples where debates about evidence shift into debates on whether or not the evidence is scientific—i.e., about the philosophy of science, and the legitimacy of evidence.
3. Demonstrate, using actual examples from historical psychology, a more sophisticated understanding of the process of scientific inquiry than the simplistic experimental model.
4. Understand the concept of a ‘paradigm’ as Kuhn originally meant it: as a ‘model’ experiment that defines an era of a discipline or a scientific culture.
5. Distinguish between observational studies, experimental interventions and modeling, and the scientific value of each.
6. Apply their understanding of paradigm cases like Newton’s *Experimentum Crucis*, and Hermann Ebbinghaus’ memory experiments as the paradigm of experimental psychology, to judge the validity of new and novel research proposals.
7. Recognize and manipulate the logical structure of scientific proposals (i.e. The H-D method), including correctly interpreting evidence gathered as bearing on the hypothesis being tested. As this is historical, the students should also recognize that the reasoning of real, historical scientists frequently fails to follow this ideal.

8. Identify potential confounding variables in basic experimental design, and suggest ways those variables may be controlled.
9. Identify the strengths and weaknesses of non-experimental social science methods, specifically observational research (both naturalistic and participant-observation) and modeling; and apply this analysis to real-word cases (Rosenhan 1973, Haney, Banks, and Zimbardo 1973, e.g.). (=
10. Demonstrate a working knowledge of the intellectual 'environment' that led to the decline of behaviorism and psychoanalysis.
11. Engage with primary sources in the history of psychology.
12. Demonstrate a working knowledge of the major distinctions in the study of the mind: psychodynamic, behavioristic and cognitive, including being able to correctly categorize major views /works, and discuss the complexity of cases on the margins (i.e. Tolman, Bandura)
13. Appreciate the social impact taxonomies and labels can have on others, and the social significance thereof.

Conceptual issues at stake

Many of the issues raised in the course of the game are appropriately philosophical. While we've instructed the students to focus on the evidence, and its relationship with the claims asserted, one is often forced to notice how few of the great conflicts in the history of psychology and psychiatry are actually about the quality of the evidence presented. The conflicts are, most frequently, actually about whether the evidence in question can be considered 'scientific.'

The APA guidelines cited above do not report *what* the key terminology, concepts and theories are—we cannot claim to cover all that a traditional Introduction to Psychology class would, but we can point out where a number of these key concepts appear:

- Conditioning: classical, operant and vicarious: Section Behaviorists subsection 4.2 on p. 142, articulated by Bandura , Harlow and Hooker. Specific experiments to be considered include Lab Emotional responses in a human child subsection 5.3 on p. GM79
- Social Psychology: Milgram and Zimbardo — two labs: Zimbardo proposes Interpersonal Dynamics in a Simulated Prison subsection 5.3 on p. GM88 to Research committee and, if time allows, Small World subsection 5.3 on p. GM91 from Milgrim.)

- Statistical concepts: Hooker's paper (Hooker 1956) sets the stage, but it will reappear in the game play of Anastasi and Spitzer. Spitzer's presentation on the Feighner Criteria should cover statistical concepts. Anastasi should present a paper about 'norms' and whether 'normal' is a statistical function of a population, or an ideal. Lab: A Re-analysis of the Reliability of Psychiatric Diagnosis subsection 5.3 on p. GM86
- Experimental Design: Research committee should have a number of distinct research designs to consider. See Section Basic scientific research section 5.1 on p. 151 in the game book as well as Section Ethics of human research section 5.2 on p. 182.
- History and Systems: Section Brief history of the concept of 'Psychology' section 2.2 on p. 23 covers most of what common textbooks cover in this section, including the rise of empirical psychology from the 19th century 'natural philosophers', early experimentalism, psychoanalysis, behaviorism and cognitivism. We do not cover 'positive psychology' or 'humanistic psychology' (I.e. Maslow).

The Philosophy of Science concepts and theories that arise are:

- The logic of paradigms ('normal' science), and the revolution periods of transition between them. (Section The Paradigmatic experiments subsection 5.1 on p. 153)
- Hypothetical-Deductive model of scientific reasoning, including verificationism and falsificationism. The Logic of evidence subsubsection 5.1 on p. 160, Research committee discussions.
- Modeling in scientific inquiry—Section Models in Science subsubsection 5.1 on p. 180. Lab: Mental rotation of three-dimensional objects subsection 5.3 on p. GM81.
- Complications of observation — Section Observational research subsubsection 5.1 on p. 163. Lab: On being sane in insane places subsection A on p. GM122.

I've noted where in the games these issues arise.

Some of these issues that will make an appearance in the game are:

Issue	Partisans	Session
-------	-----------	---------

Homosexuality <i>is / is not</i> a mental disorder	Spitzer, Socarides, Bieber, Marmor	1971 Marmor Proposal 1 on p. R136 1971 Spiegel Proposal 1 on p. R140 and Green Proposal 1 on p. R108 1971 Gittings Research 1 on p. R66 1972 Green Writing Assignment 1 on p. R108 and Spiegel Writing Assignment 1 on p. R140) Socarides Writing Assignment 1 on p. R124 and Bieber Writing Assignment 1 on p. R131 Socarides Proposal 1 on p. R123 1973 Nomenclature Symposium. Include Gold Writing Assignment 1 on p. R56
The 'medical model' <i>is / is not</i> a suitable approach to understanding the human mind.	Szasz, Albee v Spitzer and all the MDs	1972 Symposium: Szasz Writing Assignment 1 on p. R99, Marmor Writing Assignment 1 on p. R137 and Albee Writing Assignment 1 on p. R19
The mentally ill <i>must / should be treated exclusively by psychiatrists / psychologists and psychologists / exclusively psychologists</i>	Albee	1972 Hopcke Proposal 1 on p. R112 and Albee Proposal 1 on p. R19 Mediated by Fordney-Settlage Proposal 1 on p. R36 in the Role Sheet Book
Are mental health norms defined in terms of statistical frequencies of behavior, or in terms of the 'ideal' or 'proper' functioning of an individual?	Anastasi	1973 Anastasi Writing Assignment 1 on p. R25 and Spitzer Writing Assignment 1 on p. R92
Psychoanalysis <i>is / is not</i> scientific.	Psychoanalysts v. Cognitivists and Behaviorists	1972 Marmor Writing Assignment 1 on p. R137 and Szasz Research 1 on p. R99 1973 Spitzer Research 1 on p. R92 and Marmor Research 1 on p. R136
Mental disorders should be classified according to <i>observable symptoms / underlying mechanisms</i> .	Spitzer and the behaviorists v. psychoanalyst faction	1972 Spitzer Writing Assignment 1 on p. R92, Anastasi Writing Assignment 1 on p. R25 and Leif Writing Assignment 1 on p. R116
There <i>should / should not</i> be psychological study of specific groups, I.e. Women?	Fordney-Settlage, Anastasi	1972 Fordney-Settlage Research 1 on p. R36 1973 Anastasi Proposal 1 on p. R26 and Anastasi Proposal 2 on p. R27 1974 Clark - 1974 Writing Assignment 1 on p. R32 1975 Fordney-Settlage Writing Assignment 1 on p. R37

What does scientific research on the mind look like? Does it involve modeling and adductive reasoning, or is it limited to correlations between observable behaviors?	Cognitivists and Psychoanalysts v. Behaviorists	1971 Miller Research 1 on p. R177 1975 Marr Writing Assignment 1 on p. R172 and Marr Research 1 on p. R172
What are the ethical limitations on psychological / psychiatric research?		1971 Emotional responses in a human child subsection 5.3 on p. GM79 1973 Zimbardo Research 1 on p. R103 1975 Zimbardo Writing Assignment 1 on p. R104
There <i>are / are not</i> innate structures in the mind	Chomsky, Piaget, Behaviorists, Psychoanalysts	1975 Piaget Writing Assignment 1 on p. R75 and Chomsky Writing Assignment 1 on p. R169
Intellectuals—specifically social scientists— <i>have / do not have</i> a special responsibility to democratic society.	Miller, Chomsky, Clark, Tyler	1971 Miller Writing Assignment 1 on p. R177 1974 Chomsky Writing Assignment 2 on p. R169, Clark - 1974 Writing Assignment 1 on p. R32 v. Tyler Proposal 1 on p. R80
How do we define a mental illness / disorder?	Spitzer, main factions	Spitzer Proposal 4 on p. R92 Psychoanalysts Proposal 1 on p. R105 Behaviorists Proposal 1 on p. R144

Table 5.15: Major Issues for debate

Research design

I've included a 'textbook' style section (Testing causal claims experimentally subsubsection 5.1 on p. 168) of the game book that introduces the standard independent variable-dependent variable-controlled variable logic that is included in most research methods in psychology courses.

I prefer, of course, to introduce these ideas with real historical examples, as I do in Section The Paradigmatic experiments subsection 5.1 on p. 153.

I included the 'textbook' section to give the research committee a quick heuristic for judging the validity of proposed experiments. But this simplistic approach is largely misguided and the research committee may stumble as a result. For example, they will encounter these real, historical examples that simply don't fit the model:

- We know that Rosenhan measured the time spent inside a mental health institution as a dependent variable, but what was the IV?
Was there a control group?
- We know that the IV in Zimbardo's prison 'experiment' is assign-

ment to either 'prisoner' or 'guard' role, but what is his DV? Is he measuring anything?

- Milgram measured the number of hops a letter would take in his 'small world' experiment, but what is his manipulation of the IV?

We want the research committee to raise these kind of questions when these proposals come before the committee.

Readings

Psychology is a heterogeneous field, and including Psychiatry in the game makes it even more complex. This game is intentionally designed to embody the kind of interdisciplinary dialogue that occurred in the 1970's and gave rise to cognitivism and our theoretically-pluralistic understanding of sexuality and human behavior in general.

Common reading

As such, each student ought to read the readings contained on his or her role sheet, a full summary of which is included in table Overview of reading assignments, by character Table 5.17 on p. GM68.

The challenge of this game is to ensure that there is a common base among the various approaches to the mind to ensure productive discussion. I've chosen Freud's *Introductory Lectures on Psychoanalysis* because everyone—friend and foe—would have needed to understand and respond to Freud during this period.

The *Lectures* are not difficult, but they are long. I use the reading schedule laid out in Box Common reading Table 5.16 on p. GM64 over three class sessions, and focus on the points listed therein. Of course, you are welcome to adjust this to your needs.

Class 1, first class: Part 1 (Lectures 1-4): Parapraxes, with special attention paid to:

1. The topographic hypothesis and introduction of the unconscious (p. 25) (Thesis #3 in Classical subsubsection 4.1 on p. 125)
2. The economic hypothesis and libidinal energy (p. 26) (Thesis #3 in Classical subsubsection 4.1 on p. 125)
3. The idea of the appearance a 'disturbing intention' (Lecture III, pg. 63-72) and the relation between the 'disturbing intention' and the 'disturbed one' (p. 75-78)—which sets the stage for his mechanisms of representation (see Classical Table 4.1 on p. 128).
4. The Dynamic hypothesis (p. 82 and again p. 94) (Thesis #1 in Classical subsubsection 4.1 on p. 125).

Class 2: Dreams (Lectures 5-15)

1. Transformations (I.e. 'Mechanisms of representation' in Dream-work) in Lecture 11: Condensation (p. 210-214), Displacement (214-215), Plastic portrayal (215-219), Opposition (219-223). (see Classical Table 4.1 on p. 128)
2. Lecture 13: first instance of thesis #4 Genetic hypothesis: 'regressive' nature of dreams and origination in childhood
3. Instance on sex as source of all energy, including in Children (p. 258)

Class 3: Neurosis (lectures 16-28)

1. Definition of delusion (p. 310)
2. Symptoms as meaningful ("having a sense") – Lecture 17, p. 318-333, includes discussion of 'obsessional neurosis' (monomania / OCD) throughout. Good for comparison against other approaches.
3. Economic hypothesis reappearance - p. 340.
4. Lecture 14 - Resistance and Repression. Def. of repression on pg. 364, early characterization of Thesis #5 Structural Hypothesis on pg. 365, p. 425 and 437.
5. Transference neurosis (p. 371)
6. Discussion of homosexuality ('inverts') p. 377-378, use of phrase 'ego-syntonic' on pg. 436, which may or may not appear in Spitzer's proposal.
7. Restatement of economic hypothesis in connection with *pleasure principle* p. 443
8. Lecture 23: Classical psychoanalysis approach to classifying and treating symptoms – 'symptoms' defined on pg 445, definition of 'falling ill' on pg. 480.
9. Lectures 25-27 cover specific neurosis, and 28 covers treatment.

Table 5.16: Core text reading list

I have also included Bacon's *Novum Organon* and Ebbinghaus *Memory* as supplemental core texts. I enjoy both, and believe students would be better off having read them, but they are not necessary for game play.

Weekly readings

There will be primary source readings attached to each conference, depending on the actions of the Research Committee. See Section Lab Activities section B on p. GM176 for more.

Overview of assignments, by character

Paper	Proposals
G Albee	Albee, G. W. (1970) "The uncertain future of clinical psychology" <i>American Psychologist</i> 1071-1080

- A. Anastasi Anastasi, A. (1972) "The cultivation of diversity" *American Psychologist* 27(12), 1091-1099
 Anastasi. (1967). "Psychology, psychologists, and psychological testing." *American Psychologist* 22 (4), p. 297-306
- A. Bandura Bandura, A. (1965). Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 1-55). New York: Academic Press.
 Bandura, A. (1971). Vicarious and self-reinforcement processes. In R. Glaser (Ed.), *The nature of reinforcement* (pp. 228-278). New York: Academic Press. [Classic]
 Bandura, A. (1974). Behavior theory and the models of man. *American Psychologist*, 29, 859-869.
- I. Bieber Bieber, I. (1962). Homosexuality: A Psychoanalytic Study: By Irving Bieber, et. al. New York: Basic Books, New York
- N. Chomsky Chomsky, N. (1959). "Review of Skinner's Verbal Behaviour" *Language* 35: 26-58. Electronic copy of 'review' with preface: <http://cogprints.org/1148/0/chomsky.htm>
 Chomsky, N. (1967). "Recent Contributions to the Theory of Innate Ideas" *Synthese* 17: 2-11.
 Chomsky, N. (1967). "The Responsibility of Intellectuals" The New York Review of Books, Feb 23, 1967. Available at <http://www.chomsky.info/articles/19670223.htm>
- K. Clark Phillips, L. "Recontextualizing Kenneth B. Clark: An Afrocentric Perspective on the Paradoxical Legacy of a Model Psychologist-Activist" from *Evolving Perspectives on the History of Psychology*, Pinkren, W and Dewsbury, D.
 Clark, K. (1971) "The pathos of power: A psychological perspective" *American Psychologist* 26(12) 1047-1057
- D. Fordney Settlage, DS, Baroff S, Cooper D 1973. Sexual experience of younger teenage girls seeking contraceptive assistance for the first time. *Family Planning Perspectives*, Vol. 5, No. 4 Autumn, 1973, p 223-226 (available on Jstor)
 Fordney-Settlage DS 1975. Heterosexual dysfunction: Evaluation of treatment procedures. *Archives of Sexual Behavior*, Volume 4, Number 4 / July, 1975 (DOI: [10.1007/BF01541722](https://doi.org/10.1007/BF01541722))
- J. Fryer Scasta, D. (2002) "John E. Fryer, MD, and the Dr. H. Anonymous Episode" *Journal of Gay and Lesbian Psychotherapy* 6(4) 73-84
- P. Gebhard Cornbach, L. J. (1956) "The Two Disciplines of Scientific Psychology" *American Psychologist*, 12: 671-684. Available at <http://psychclassics.yorku.ca/Cronbach/Disciplines/>

- R. Gold Gold, R. "Three Stories of Gay Liberation"
<http://www.queerstories.org/custom.html>
 Moor, Paul (2001) The view from Irving Bieber's couch: "Heads I win, tails you lose." *J. Gay and Lesbian Psychotherapy*, 5(3/4):25-36. (also reprinted Sexual Conversion Therapy (Ariel Shidlo, Michael Schroeder (PsyD.), Michael Schroeder, Jack Drescher)
- R. Green Green, R. & Money, J. (1961) "Effeminacy in prepubertal boys; Summary of eleven cases and recommendations for case management." *Pediatrics* 27(2), p. 286-291
 Green, R. (1972) "Homosexuality as a Mental Illness" *International Journal of Psychiatry* 10:77-98
 Fox, K. "Vancouver – The Richard Green Interview"
http://web.archive.org/web/20030609051314/http://www.rfts.a.se/rich_greene.html, accessed 2/17/2011
- H. Harlow Harlow, H. (1958) "The nature of love" *American Psychologist* 13(12) 673-685 On Psych Classics:
<http://psychclassics.yorku.ca/Harlow/love.htm>.
 Harlow, Harry F. (1975). "Lust, latency and love: Simian secrets of successful sex." *Journal of Sex Research* 11(2): 79-90. available on jstor.org
- E. Hooker Hooker, E. (1956). A preliminary analysis of group behavior of homosexuals. *Journal of Psychology*, 42, 217-225.
 Hooker, E. (1957). The adjustment of the male overt homosexual. *Journal of Projective Techniques*, 21, 18-31.
 Hooker, E. (1958). Male homosexuality in the Rorschach. *Journal of Projective Techniques*, 23, 278-281.
 Hooker, E. (1961). Homosexuality: Summary of studies. In E. M. Duvall & S. M. Duvall (Eds.), *Sex ways in fact and faith*. New York: Association Press.
- R. Hopcke Hopcke, R. (1988) "Jung and Homosexuality: A Clearer Vision" *Journal of Analytical Psychology* 33, p. 65-80
- B. Gittings Story of Barbara Gittings:
<http://www.rainbowhistory.org/gittings.htm>
 Siegelman, M. (1972). "Adjustment of Homosexual and Heterosexual Women." *British Journal of Psychiatry* 120, 477-81.
- F. Kameny Read the legal and personal documents of Frank Kameny at
<http://www.kamenypapers.org/index.htm>
 Read the story of the 1971 APA in DC:
<http://www.rainbowhistory.org/html/apazap.htm>
- H. Lief Leif, H. (1964) "The Psychological Basis of Medical Practice" *Postgrad Med J* 40:355 doi:10.1136/pgmj.40.464.355-a
- J. Marmor Marmor, J. (1970) "The Limits of Free Association" *Arch. Gen. Psychiat.* 22, p. 160-165

- D. Marr
Marmor, J. (Ed.) (1965). *Sexual inversion: The multiple roots of homosexuality* (pp. 83107). New York: Basic Books.
- S. Milgram
Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* MIT Press (esp. Ch 1)
- G. Miller
Milgram, S. (1963) "Behavioral study of obedience." *Journal of Abnormal and Social Psychology*, 67, 371-378. Available at http://www.wadsworth.com/psychology_d/templates/student_resources/0155060678_rathus/ps/ps01.html with permission from Alexandra Milgram.
- J. Piaget
Milgram, S. (1967). "The Small World Problem" *Psychology Today*, 2 p. 60 - 67
- L. Tyler
Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." *Psychological Review*, 63, 81-97.
- C. Socarides
Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist*, 24(12), 1063-1075.
- J. Spiegel
Piatelli-Palmarini, M. (Ed.) (1980). *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, London: Routledge & Kegan Paul . Reviewed at: <http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=5843312>
- R. Spitzer
Tyler, L.E. (1969). *The Work of the Counselor*, Prentice Hall
- Tyler, L. E. (1973) "Design for a hopeful psychology" *American Psychologist*, 28(12) 1021-1029
- Socarides, C.W. (1963). "Homosexuality: A Psychoanalytic Study: By Irving Bieber, et. al. New York: Basic Books Inc., 1962 385 pp." *Psychoanalytic Quarterly*, 32: 111-114
- Socarides, C. W. (1968). *The Overt Homosexual*. Jason Aronson, Inc.
- Socarides, C.W. (1970). "Homosexuality and Medicine" *Journal of the American Medical Association* 212 (7): 1199
- Grinker, R. & Spiegel, J. P. (1945) *Men Under Stress* American Psychological Association. Was once available print-on-demand, but is not currently: www.apa.org/pubs/books/4320127.aspx
- Spitzer, R. L. and Endicott, J. "Medical and Mental Disorder: Proposed Definition and Criteria" in *Critical Issues in Psychiatric Diagnosis*, Spitzer and Klein, e.d. See also Klein, D. "A Proposed Definition of Mental Illness", in the same volume.
- Spitzer. Research diagnostic criteria (RDC). Biometrics research (1975) pp. 34 (available at <http://www.garfield.library.upenn.edu/classics1989/A1989U309700001.pdf>)

- T. Szasz Spitzer, R. and Fleiss, J. (1974). "A Re-analysis of the Reliability of Psychiatric Diagnosis" *Brit. J. Psychiat.* 125, 341-7
- T. Szasz Popper, K. R. (1959). *The Logic of Scientific Discovery*. New York: Basic Books
- Szasz, T. (1960). "The Myth of Mental Illness" *American Psychologist*, 15, p. 113-118 available at <http://psychclassics.yorku.ca/Szasz/myth.htm>
- P. Zimbardo Haney, C. Banks, C. & Zimbardo, P. (1973). "Interpersonal Dynamics in a Simulated Prison" *International Journal of Criminology and Penology* 1, p. 69-97
- Zimbardo, P. G. (1972). *The Stanford Prison Experiment a Simulation Study of the Psychology of Imprisonment*. Philip G. Zimbardo, Inc.

Table 5.17: Overview of reading assignments, by character

*Overview of events and assignments in the game***Overview of assignments, by character**

Paper	Proposals	Research
G Albee—Nomenclature Committee, expiring '72 —Board member 'At large B' '71-'72* (1)	Parallel medical system for clinical psychology '72	'Medical Model' – symposium '72 1
A. Anastasi—Program '71-'73 (1)	Proposal to fix gender-specific language in guidelines (1), Women in Psych (1) and Psych of Women '74 (2)	
"On psychometrics" - with Spitzer's 'Feighner criteria' (1)		
A. Bandura—VP '71 1 (Board '71-'73)		
"Behavioral theory and models of man" '72 1	'73 – reject Bieber	
I. Bieber		
Comment on Green / debate Marmor '72, Report on sexual deviation (Freudian) '73 1, Symposium on nomenclature '73	Oppose Marmor	Gold's '71 proposal to remove 302.0 from DSM-II, Propose taskforce '71, report '73
N. Chomsky		
"The responsibilities of Intellectuals" - symposium '74 (2), Debate with Piaget '75 1	Oppose L. Tyler principle '73	

K. Clark—Research '71-'73 1

"psychology as a force of social change" - symposium '74 (1)

Oppose L. Tyler principle '73,
Proposal to create affirmative action plan '75

D. Fordney Settlage*

Paper critiquing the focus on male sexuality '75 (1)

Propose balance compromise
Albee and Hopcke on psych.
Practicing in hospitals, but not prescribing drugs.

J. Fryer

"I am a homosexual, and a psychiatrist." - symposium '71 1

Proposal to deem the use of aversion therapy as immoral '74 (See 5 - association business)

K. Freund

P. Gebhard— Program '71-'74

Summary of the Kinsey report '72 1

Honor of E. Hooker '72 (1)

R. Gold*

Symposium on Nomenclature '73 (1)

Support Marmor's '71 proposal

Draft proposal with Spitzer '73

R. Green*

Report on history of homosexuality '72 (1) (with J. Spiegel), Task force symposium on Nomenclature '74

Taskforce on the history of homosexuality in Psych '71 (with Spiegel)

Propose "Bringing the Clinic into the Undergraduate Classroom" (B)* 1

H. Harlow—Board '71 1

"Lust, latency and love" (1) '72

Oppose Fryer's proposal to deem aversion therapy immoral '74

Propose "Operant Conditioning: Role in Human Behavior" B 1

E. Hooker—Program 1 '71-'72

"The mental health of non-patient male homosexuals" - symposium '71, (1)

Oppose dissolution of Nomenclature committee.

R. Hopcke*

Jungian approaches to homosexuality '75 1

(1) '72

B. Gittings*

Symposium on your experiences, '71	Organize and Chair the 'big' symposium "Should 'homosexuality' be in the DSM-III?" in '73 where Green, Marmor, Bieber, Gold, Socarides, Spitzer should speak.	Propose Siegelman "Adjustment of homosexual women" 1971 1
F. Kameny*		
"Gay, Proud and Healthy" - symposium '71 (1)	Organize and Chair the 'big' symposium "Should 'homosexuality' be in the DSM-III?" in '73 where Green, Marmor, Bieber, Gold, Socarides, Spitzer should speak.	
H. Lief		
Paper on genetic hypothesis and diagnostic taxa '74 (1)	Oppose Spitzer's nomenclature proposal to write the DSMIII without the genetic hypothesis '72	
J. Marmor—Research '71-'74		
"Limitations of Free Association" - Symposium on the 'Medical Model' '72, 1, Comment on Green / debate Bieber '72, Symposium on Nomenclature '73	Propose 'homosexuality' be removed from category 302. (1) '71 Propose a condemnation of Socarides' JAMA paper. (2) '73	Propose 'To Sleep, Perchance to Dream'— B (1)
D. Marr*		
'Levels' of explanation paper '75		Propose 'Information Processing Capacity: A Visual Demonstration of the Magical Number Seven' — (B) '72 (1)
S. Milgram*—Board member 'A'* (1)		
Summary of obedience experiments / findings '74 (1)	Oppose Clark's affirmative action plan - '75	Propose 'small world' experiment '74 1
G. Miller—Board '71-'72 (1)		
"Psychology as a means of promoting human welfare" (1) '71		Present Mental Rotation '71 (Shepard / class data) (1)
J. Piaget*		

Debate with Chomsky '75 1		
L. Tyler—Research '71-'72 1		
"Design for a hopeful psychology" -1 '73	Leona Tyler principle (1) '73	
C. Socarides		
Comment on Green / debate Marmor '72, Report on sexual deviation (Freudian) '73 1, Symposium on nomenclature '73	Oppose Marmor and Gold's '71 proposal to remove 302.0 from DSM-II	
J. Spiegel—Nomenclature '71-'74 1		
Report on history of homosexuality '72 (with R. Green) 1 If elected, Presidential address '75	Propose, with Green, a task force on the history of the treatment of homosexuality (1) '71.	Defense Mechanism 1
R. Spitzer—Nomenclature '71-'76 1		
"The Feighner Criteria" 1'72, Nomenclature / Task Force sponsored symposium on homosexuality in '73, With the committee / task force, open symposium in '74 on mental illness	As Nomenclature committee: proposal write DSM according to description of symptoms, not causes (2 '72, Proposal to disband Nomenclature and form Spitzer Taskforce (3 '73, With R. Gold, draft proposal to remove 'homosexuality' immediately in '73 (1), Propose Taskforce definition of mental illness (4) '75	Meta-analysis of reliability of psychiatric diagnosis (1 '73
T. Szasz		
As part of the symposium (with Albee) on the "Medical Model", present: "The myth of mental illness," 1 '72	Oppose committee on psychology of women '73 (on grounds of reification), Propose there is no definition of mental illness (1) '75	
P. Zimbardo*		
Paper on the ethical problems of the Stanford Prison Study, and any holes in the 1968 APA guidelines. 1 '75		Propose Zimbardo '73 (1) (if character is not used, the proposal should come from the gamemaster)

Table 5.18: Overview of specific assignments, by character

Papers in italics are conditional. If not elected, the students should present them as papers.

Research items in italics are NOT included in the character sheet, and can be moved around by the instructor.

Not included in games smaller than 26.

Overview of topics

See Major Issues for debate Table 5.15 on p. GM62.

Issue / Year	Proposer	Objections	Mediator	Related papers
Demedicalization (1971)	Marmor	Socarides and Bieber	Spiegel and Green	
Limitation of care to hospital settings and psychiatrists (1972)	Hopcke	Albee		Albee Writing Assignment 1 on p. R19
Psychology of Women (1973)	Anastasi	Szasz		Fordney-Settlage Writing As- signment 1 on p. R37
Leona Tyler princi- ple (1973)	Tyler	Chomsky, Clark		Chomsky Writing Assignment 1 on p. R169
DSMIII according to symptoms not causes (1973)	Spitzer / Nomenclature	Leif		Spitzer Writing Assignment 1 on p. R92, Leif Writing Assign- ment 1 on p. R116
Ban Aversion Therapy in the treatment of ho- mosexuality (1974)	Freund	Harlow		
Affirmative Action (1975)	Clark	Milgram		Clark - 1974 Writing Assign- ment 1 on p. R32, Clark - 1974 Writing Assignment 2 on p. R32

Table 5.19: Overview by topic

Overview of elections, by year

Retiring members in parenthesis

Year	VP	Board at large	Research (replacing)	Nomen- clature (replacing)	Program (replacing)
1971	1972: (Ban- dura)	A (Milgram)			
1972	1973	B (Albee)	(Tyler)	(Albee)	(Hooker)

1973	1974		(Clark)		(Anastasi)
1974	1975	A	(Marmor)	(Spiegel)	(Gebhard)
1975	1976	B	(elected 1972)		(elected 1972)

Table 5.20: Elections to be held each year

Initial credibility of each character

Before the game begins, the credibility of each character ought to be established. The values displayed in table Initial credibility Table 5.21 on p. GM74 are on the top of each individual's game sheet. It is important, however, for everyone in the game to know how much credibility the various players have. To that end, it is advisable for the game master to start the game by publicly distributing the actual 'credibility tokens' (Credibility Points subsection 5.2 on p. GM51) during the start of the first class, announcing who has what credibility. Ron Gold, because of his position as a journalist for a well-known and respected magazine, starts the game with 10 credibility points that he *must* distribute before the game begins. His own credibility, as a non-academic is 'o.'

Name	Initial Credibility
Robert Spitzer	5
George Miller	10
Anthony Bandura	10
Harry Harlow	10
Noam Chomsky	15
Leona Tyler	7
Anne Anastasi	7
John P. Spiegel	5
Evelyn Hooker	14
George Albee	5
Ken Clark	10
Judd Marmor	6
Thomas Szasz	-10
Irving Beiber	4
Charles Socarides	4
John Fryer	2
Kurt Freund	2
Paul Gebhard	4
H. Lief	2
Robert Hopcke	0

Jean Piaget	15
Dr. Fordney Settlage	7
Kameny	0
Gittings	0
Marr	0
Gold	0
Milgram	15
Zimbardo	0

Table 5.21: Initial credibility

Committees, Publications and Exhibits

Likely actions indicated in

ALLCAPS

year	Board of Directors	Research Committee	Nomenclature Committee	Program Committee
Initial mem-ber-ship	Harlow (1971) Miller (1972) Bandura (1973) [A: Milgram]* [B: Albee]*	L. Tyler (1972) K. Clark (1973) J. Marmor (1974)	G. Albee (1972) J Spiegel (1974) R Spitzer (1976)	E. Hooker (1972) A. Anastasi (1973) P. Gebhard (1974)
1971	Proclamations removing Homosexuality heard Marmor Proposal 1 on p. R136 and <i>sent to Nomenclature</i>	Miller reporting Mental Rotation Miller Research 1 on p. R177.	Accept charge to consider removing 302.0	Prepare Dallas Conference schedule, distribute at the end of this conference
	APPROVE taskforces for Green (Green Proposal 1 on p. R108) and Bieber (Bieber Proposal 1 on p. R131)	APPROVE Gittings Research 1 on p. R66 REJECT Emotional responses in a human child subsection 5.3 on p. GM79		

1972	Recognize Hooker Gebhard Proposal 1 on p. R50 Limit drugs to psychiatric, but recognize psychology access in hospital Hopcke Proposal 1 on p. R112 and Albee Proposal 1 on p. R19	APPROVE Fordney-Settlage Research 1 on p. R36	Create proposal for observation, not theory-based DSM (Spitzer Proposal 2 on p. R92)	Prepare Honolulu Conference schedule, distribute at the end of this conference
1973	Consider condemnation of Socarides Marmor Proposal 2 on p. R136 Spitzer's proposals Spitzer Proposal 1 on p. R91 and Spitzer Proposal 3 on p. R92	APPROVE Zimbardo Research 1 on p. R103 Spitzer Research 1 on p. R92	Dissolution in favor of Spitzer task force	Prepare Philadelphia Conference schedule, distribute at the end of this conference
	SEND Spitzer Proposal 1 on p. R91 to vote of the membership			
	Leona Tyler Principle			
	Tyler Proposal 1 on p. R80			
1974	Consider Psy. Of Women (Anastasi Proposal 2 on p. R27), Women in Psy. (Anastasi Proposal 1 on p. R26)	APPROVE Milgram Research 1 on p. R70 Fryer Research 1 on p. R41	DISSOLVED—or hold open session on def. of 'mental illness'	Prepare Chicago Conference schedule, distribute at the end of this conference
	BAR Aversion therapy Freund Proposal 1 on p. R47			
1975	Consider Affirmative Action plan Clark Proposal 1 on p. R32	Propose rewritten Ethics guidelines Zimbardo Writing Assignment 1 on p. R104	<i>dissolved</i>	No actions

Table 5.22: Likely Actions by Committee

Labs and possible research proposals

Various roles are tasked with proposing research studies to the research committee, conducting that research if it is chosen, and presenting the findings at the following conference. Not everything proposed can be carried out in an undergraduate classroom. When the proposal is impossible, the instructor should distribute the actual paper for evaluation *as if* it had been carried out. The instructors is, of course, encouraged to use classroom activities with which he or she is familiar.

I've pulled a number of experiments from the APA's *Activities Handbook for the Teaching of Psychology* v. 1–4 that correspond to many of the topics the class will be discussing. They are listed here, and attached as PDFs at the end of this document, if you do not have access to the *Handbooks*. The characters listed here are *suggestions*. The instructor should distribute these as he or she sees fit.

Title, author	Location in APA Handbooks	Suggested character, conference	Appendix
Accuracy of Observation, Paul J. Woods	v. 1,#2	Tyler or Anastasi 1971	Appendix Accuracy of Observation subsection B on p. GM176 of the Instructor's Manual
Operant Conditioning: Role in Human Behavior, Edward Stork	v. 1,#23	Harlow	Appendix Operant Conditioning: Role in Human Behavior subsection B on p. GM179
Operant Conditioning Demonstration, Patricia Keith-Speigel	v. 1,#24	Harlow or Hooker	Appendix Operant Conditioning Demonstration subsection B on p. GM181 of the Instructor's Manual
Defense Mechanisms, Jack J Greider	v. 1,#75	research- task:spiegel	Defense Mechanisms subsection B on p. GM184 of the Instructor's Manual
To Sleep, Perchance to Dream, Judy T. Benjamin, Jr.	v. 1,#80	Marmor Research 1 on p. R136	Appendix To Sleep, Perchance to Dream subsection B on p. GM186 of the Instructor's Manual
Mental Illness, James M. Gardner	v. 1,#81	Fryer Research 1 on p. R41	Appendix Mental Illness subsection B on p. GM190 of the Instructor's Manual

JAWS: Demonstrating Classical Conditioning, Randolph A Smith	v. 2,#19	Harlow or Hooker	Appendix Classical Conditioning subsection B on p. GM193 of the Instructor's Manual
Human Operant Conditioning, John K. Bare	v. 2,#20	Harlow Research 1 on p. R151	Appendix Human Operant Conditioning subsection B on p. GM196 of the Instructor's Manual
Bringing the Clinic Into the Undergraduate Classroom, David M. Young	v. 3,#45	Green Research 1 on p. R108, 1974-1975	Appendix Bringing the Clinic into the Undergraduate Classroom subsection B on p. GM199 of the Instructor's Manual
Discovering the Relationship Between Operational Definitions and Interobserver Reliability, Angela H. Becker	v. 4,#15	Could be instead of Spitzer, 1973, or in combination	Appendix Discovering the Relationship Between Operational Definitions and Interobserver Reliability subsection B on p. GM207 of the Instructor's Manual
Information Processing Capacity: A Visual Demonstration of the Magical Number Seven, Fairfid M. Caudle	v. 4,#43	Marr Research 1 on p. R172	Appendix Information Processing Capacity: A Visual Demonstration of the Magical Number Seven subsection B on p. GM215 of the Instructor's Manual
The Role of Prior Information in Dream Analysis, Douglas A. Bernstein	v. 4,#80	Marmor Research 1 on p. R136, 1973	Appendix The Role of Prior Information in Dream Analysis subsection B on p. GM221 of the Instructor's Manual

Table 5.23: Possible lab activities

Historical studies

When and if the Research Committee approves experiments that are based on historical realities, the game master ought to distribute the primary source report of said experiment for discussion in the following week. These are available under 'Primary Source Materials' on the GitHub site, and included in the Appendix.

- 1971 Lab: Emotional responses in a human child subsection 5.3 on p. GM79
- 1971: Lab: Mental rotation of three-dimensional objects subsection 5.3 on p. GM81.

- Shepard and Metzler 1971.
- 1972: Adjustment of Homosexual and Heterosexual Women subsection A on p. GM118
 - Siegelman 1972
- 1973: Lab: On being sane in insane places subsection A on p. GM122.
 - Rosenhan 1973
- 1974: Lab: Interpersonal Dynamics in a Simulated Prison subsection 5.3 on p. GM88.
 - Haney, Banks, and Zimbardo 1973
- 1974: Lab: A Re-analysis of the Reliability of Psychiatric Diagnosis subsection 5.3 on p. GM86.
 - Spitzer and Fleiss 1974
- 1975: Lab: Small World subsection 5.3 on p. GM91
 - Milgram 1967

5.3 *Research Committee Proposals*

These are provided to the instructor in case the students fail to come through on research design. They can either be handed to the relevant student to inspire creativity, or “submitted” to the research committee by a mysterious non-player character.

*Emotional responses in a human child.***Background:**

Many psychologists—Freud included—have held, with little evidence in support, that the human mind is built on a variety of instincts, such as self-preservation, sexual activity, etc. Emotional responses to fearful stimuli, such as rats and spiders, is commonly considered to be innate, possibly as a result of evolutionary pressures to avoid infectious or poisonous creatures. This experiment seeks to condition a fear-response in a human infant, thus establishing that there is no need for theoretical innate entities in our explanation of emotion in humans.

Rationale:

The success of Pavlov's work conditioning reflex responses to novel stimuli in canines has shown that behaviors previously believed to be instinctual are likely to have resulted from conditioning. The current experiment seeks to determine if that insight can be extended to humans, by conditioning an emotional response in a child—fear—that is widely believed to be instinctual. Demonstrating that these reflexes can appear without appeal to 'instinct' or 'adaptation' undermines the need for theoretically innate entities in our understanding of human behavior.

Experimental Design:

A child of a single destitute mother, currently employed as a wet nurse at a local hospital (Hopkins), has been recruited for the experiment. Given the nature of his mother's employment, the child is familiar with the clinical setting of a hospital, and hence an ideal subject for this experimental protocol. Previous work with this subject at the age of 8 months has demonstrated that he exhibits fear-like responses to loud sounds: the experimenter stood behind the subject, outside of eye sight and struck a steel bar with a hammer. On the first presentation, the subject startled and raised his hands. On the second, he began to tremble. On the third, he cried and seemed to be having a fit.

The experimenter proposes the following:

At 9 months of age, we will present the subject with (randomly): a white rat, a rabbit, a dog, a monkey, with masks with and without hair, cotton wool, and burning newspapers. Given the child's upbringing, it is unlikely that he has ever encountered any of these objects before. We expect him to have no emotional response to any of them, but if he responds, that object will be removed from the study before proceeding.

At the age of 11 months, one of the objects will be presented to him again. When he moves towards contact with the object, the experimenter will make the loud noise already established as causing a fear reaction. Each movement towards the stimulus object will cause another loud noise stimulus. Once the fear reaction is well conditioned, the experimenter will present the subject with the fear-conditioned object and record the response. The child's reaction than will be compared with other similarly-aged children's reaction for typicality of fear-reactions in children.

The fear-conditioned object will be reintroduced over the subsequent weeks and months at regular intervals to determine the persistence of the conditioned response.

Significance and Contribution

This research has the potential to experimentally confirm or deny the commonly held belief that the fear response is innate, or at least instinctual. Pavlov's experiments with classical conditioning in dogs has shown that reflexes that were previously believed to be instinctual—such as salivation—could be conditioned in response to novel stimuli. If Pavlov's approach is to be applied to humans, it is incumbent that psychologists determine the existence and nature of human instinctual reactions, and if they can be conditioned like those of the canine.

References

- Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.*

Mental rotation of three-dimensional objects

Background:

Mentally rotating 3-dimensional objects is one indicator of spatial reasoning in humans. And it is one that is tempting to explain in terms of internal mental imagery. The 'cognitive hypothesis' holds that in order to solve mental rotation problems, a cognitive representation of the presented object must be rotated in the mind before an identification can be made - and it is that hypothesis that we wish to test here.

Rationale:

When asked to match three-dimensional objects presented in two-dimensional format, individuals report imagining the object in three-dimensions and rotating them mentally to test against the other possibilities. Introspective reports are, of course, notoriously difficult to address in a scientific way, but that does not mean that we cannot study the phenomenon. If individuals are manipulating mental representations, we would expect a measurable difference in reaction time given a matching task.

Experimental Design:

A number of adult subjects (the size of the class) will be presented with pairs of two-dimensional line drawings of three-dimensional blocks. Each subject will be asked to indicate as quickly as possible if the two were drawings of the same object rotated in space or different objects. Half of the experimental set will be rotated versions of the same object, the other half not. They will be presented to the subjects in random order. The time it takes to respond will be measured using computer software. We hypothesize that the amount of time necessary to respond will be correlated with the angle of rotation of the two objects, thus establishing the cognitive hypothesis.

The materials are available on the APA website: '<http://opl.apa.org/Experiments/About/AboutMentalRotation.aspx> Your gamemaster will need to set up an account here: <http://opl.apa.org/Main.aspx>.

Significance and Contribution

This research has the potential to discover observable data that is consistent with the introspective reports of individuals.

On Being Sane in Insane Places

Background

However much mental health practitioners may be personally convinced that we can tell the normal from the abnormal, the evidence is simply not compelling. It is commonplace, for example, to read about murder trials wherein eminent psychiatrists for the defense are contradicted by equally eminent psychiatrists for the prosecution on the matter of the defendant's sanity. More generally, there are conflicting datasets on the reliability, utility, and meaning of such terms as "sanity," "insanity," "mental illness" and "schizophrenia." Finally, as early as 1934, Ruth Benedict suggested that normality and abnormality are not universal. What is viewed as normal in one culture may be seen as quite aberrant in another.

Rationale

How do we know precisely what constitutes "normality" or mental illness? Conventional wisdom suggests that specially trained professionals have the ability to make reliably accurate diagnoses. In this research, however, we intend to challenge that assumption. What is—and what is not—"normal" may have to do with the labels that are applied to people in particular settings.

Experimental Design

Eight sane people, of varied backgrounds, will gain secret admission to 12 different hospitals, from varied geographical regions of the United States. Their diagnostic experiences will constitute the data of the study. The 'pseudopatients' will call the hospital for an appointment, complaining of 'hearing voices.' When asked what the voices said, they will reply that it was unclear, but that they were 'empty,' 'hollow,' and 'thud.' The pseudopatients will report that voices were unfamiliar and in the same gender as the pseudopatient. After admission to the psychiatric ward, the pseudopatient will cease simulating symptoms of abnormality.

The amount of time it takes for the pseudopatient to be released, along with any diagnoses and treatments, will be recorded.

Significance and Contributions

This study provides an opportunity to test the reliability of psychiatric diagnosis in the 'real world,' rather than the controlled environment of a university lab. The hospital environment imposes a special environment on its members in which the actions of a normal person

could be misinterpreted as 'insane' or 'abnormal.' This study will determine the extent of that influence on psychiatric diagnosis.

Works Cited

R. Benedict, J. Gen. Psychol., 10 (1934), 59.

Adjustment of Homosexual and Heterosexual Women

Background

The traditional psychiatric belief that homosexual men are emotionally unstable has recently been challenged by Evelyn Hooker's study of non-prisoner non-patient homosexual men. There have been a few similar studies on women. The contention that homosexual women are neurotic has typically been voiced by clinicians reporting on their own patients. One exception is the recent psychometric investigation by Kenyon (1968) who studied a non-clinical group of English homosexual women, and concluded that they were higher in neuroticism than a comparison group of heterosexuals. In contrast to the 'illness' notion of homosexuality, the authors of three psychometric studies dealing with non-clinical homosexuals and heterosexuals reported that heterosexual women were not better adjusted than homosexuals. (Armon, 1960; Freedman, 1968;)

Rationale

The paucity of research in this area is exemplified by the fact that a total of only four studies, noted above, have been found to date. Even the clinical literature, which is replete with case studies and therapeutic discussions concerning male homosexuality, is strikingly sparse in the area of lesbianism. The present study is proposed to add to the small body of data we now have on the adjustment of homosexual versus heterosexual women.

Experimental Design

Working with the leadership of the New York branch of the Daughters of Bilitis, a questionnaire will be sent out to recruit members for the study. An additional questionnaire will be distributed through a popular homophile bookstore in Greenwich Village, New York. And an equivalent number of heterosexual women will be recruited from the undergraduate and graduate population of local colleges and universities.

Several different instruments will be used to measure the overall psychological adjustment of the participants, including Scheier and Cattal's Neuroticism Scale Questionnaire (NSQ) (1961) tests of Alienation and Trust (Struening & Richardson, 1965), Goal Directedness, Self-Acceptance and Sense of Self (Dignan, 1965), Dependency (Comry, 1964), Nurturance (Harvey et al. 1966) and Neuroticism (MacGuire 1966). The Crowne and Marlow Social Desirability Scale (1960) will also be used. The differences on these measures between the homosexual and heterosexual women will be compared to test the

'illness' model of homosexuality in women.

Significance and Contribution

Recent interest in and discussions of the 'illness' model of male homosexuality have almost completely ignored the parallel issues for homosexual women. This study is a small step towards closing that gap.

Works Cited.

- Armon, V. (1960) "Some personality variables in overt female homosexuality." *Journal of Protective Techniques*, 24, 292–309
- Freedman, M.J. (1968) "Homosexuality among women and psychological adjustment." *Dissertation Abstracts*, 28, 4294B.
- Hopkins, J.H. (1969). "The lesbian personality." *British Journal of Psychiatry*, 115, 1433–6
- Kenyon, F.E. (1968). "Studies in female homosexuality: IV. Social and psychiatric aspects." *British Journal of Psychiatry*, 114, 1337–50.
- Schaeier, I. H. & Cattell, R.B. (1961). *The Neuroticism Scale Questionnaire*, Champaign, Ill: Institute for Personality and Ability Testing
- Struening, E.I. & Richardson, A.H. (1965). "A factor analytic exploration of the alienation, anomia and authoritarian domain." *American Sociological Review*, 30, 768–78.
- Dignan, M.H. (1965). "Ego identity and maternal identification." *Journal of Personality and Social Psychology* 1, 476–83
- Comry, A.L. (1964). "Personality factors compulsion, dependence hostility and neuroticism." *Educational and Psychological Measurement*, 24, 75–84
- Harvey, O.J., Prather, M.S. White, J. and Alter, R.D. (1966). "Teachers' belief systems and preschool atmospheres." *Journal of Educational Psychology* 57, 373–81
- McGuire, R.G. (1966). "An inquiry into attitudes and value systems of a minority group." Unpublished Doctoral Dissertation, NYU
- Crowne, D.P. & Marlowe, D. (1960). "A new scale of social desirability independent of psychopathology." *Journal of Consulting Psychology*, 24, 349–54

A Re-analysis of the Reliability of Psychiatric Diagnosis

Introduction

Classification systems such as diagnosis have two primary properties, reliability and validity. Reliability refers to the consistency with which subjects are classified; validity, to the utility of the system for its various purposes. In the case of psychiatric diagnosis, the purposes of the classification system are communication about clinical features, aetiology, course of illness and treatment. A necessary constraint on the validity of the system is its reliability. There is no guarantee that a reliable system is valid, but assuredly an unreliable system must be invalid.

Background

Studies of the reliability of psychiatric diagnosis provide information on the upper limits of its validity. This study will consider some of the difficulties in appraising diagnostic reliability, offers a re-analysis of the available data from the literature, and suggests a possible course of action to improve psychiatric diagnosis.

Rationale

Zubin (1967) reviewed the major studies of reliability of psychiatric diagnosis performed until 1966. He noted that diagnostic reliability is referred to in three different ways: agreement between independent diagnosticians examining the same patients, stability in diagnosis over time, and similarity in diagnostic frequencies for comparable samples. It is the first sense—interjudge agreement—that is fundamental.

Recent studies of interjudge agreement of psychiatric (Schmidt and Fonda, 1956; Kreitman, 1961; Beck et al., 1962; Sandifer et al., 1964) report on agreement as to the presence or absence of a diagnosis, but they neglect to consider the rate at which diagnoses are made.

Cohen (1960) has recently developed a statistical measure (called 'kappa') of interjudge agreement that incorporates a correction for chance agreement. This study proposes to recalculate the reliability of psychiatric diagnosis from these studies based on Cohen's Kappa.

Experimental Design

The study will use existing data, culled from five recent papers measuring the interjudge agreement of psychiatric diagnosis.

Significance and Contributions

There is little doubt that the reliability of psychiatric diagnosis is being questioned at this time. If the presumed agreement of previous work depends merely on the rate of chance agreement, psychiatry must reevaluate its classification system immediately.

Works Cited

- Cohen, J. (1960) "A coefficient of agreement for nominal scales." *Educ. Psychol. Measmt.*, 20, 37–46.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J.E., & Erbaugh, J.K. (1962). "Reliability of psychiatric diagnosis: 2. A study of consistency of clinical judgments and ratings." *Amer. J. Psychiat.*, 119, 351–7
- Sandifer, M.G., Pettus, C. & Quade, D. (1964). "A study of psychiatric diagnosis." *J nerv. ment. Dis.*, 139, 350–6
- Kreitman, N. (1961) "The reliability of psychiatric diagnosis." *J. ment. Sci.* 107, 876–86.
- Schmidt, H.). & Fonda, C.P. (1956). "The reliability of Psychiatric Diagnosis: A new look." *J. abnor. soc., Psychol.* 52, 262–7
- Zubin J. (1967) "Classification of the behavior disorders." in Annual Review of Psychology (eds. P. R. Farnsworth & O. McNemar). Palo Alto, California, Annual Reviews, pp. 373–406.

Interpersonal Dynamics in a Simulated Prison

Background

After spending four years in a Siberian prison, the great Russian novelist Dostoevsky commented, surprisingly, that his time in prison had created in him a deep optimism about the ultimate future of mankind because, as he put it, if man could survive the horrors of prison life he must surely be a “creature who could withstand anything.” In the century which has passed since Dostoevsky’s imprisonment, little has changed to render the main thrust of his statement less relevant. Although we have passed through periods of enlightened humanitarian reform, in which physical conditions within prisons have improved somewhat and the rhetoric of rehabilitation has replaced the language of punitive incarceration., the social institution of prison has continued to fail. On purely pragmatic grounds, there is substantial evidence that prisons neither “rehabilitate” nor act as deterrent to future crime—in America, recidivism rates upwards of 75% speak quite decisively to these criteria. On humanitarian grounds, as well, prisons have failed: our mass media are increasingly filled with accounts of atrocities committed daily, man against man, in reaction to the penal system or in the name of it. The prison undeniably creates, almost to the point of cliché, an intense hatred and disrespect in most inmates for the authority and the established order of society into which they will eventually return. And the toll which it takes on the deterioration of human spirit for those who must administer it as well as for those upon whom it is inflicted, is incalculable.

Rationale

Attempts to provide an explanation for the deplorable condition of our penal system and its dehumanizing effects upon prisoners and guards often focus upon what might be called the *dispositional hypothesis*. While this explanation is rarely expressed explicitly, it is central to a prevalent non-conscious ideology: that the state of the social institution of the prison is due to the “nature” of the people who administer it, or the “nature” of the people who populate it, or both. That is, a major contributing cause to despicable conditions, violence, brutality, dehumanization and degradation existing with any prison can be traced to some innate or acquired characteristic of the correctional and inmate population.

The dispositional hypothesis has been embraced by the proponents of the prison *status quo* (blaming conditions on the evil in the prisoners), as well as by its critics (attributing the evil to guards and staff with their evil motives and deficient personality structures). A critical

evaluation of the dispositional hypothesis cannot be made directly through observation in existing prison settings, since such naturalistic observation necessarily confounds the acute effects of the environment with the chronic characteristics of the inmate and guard populations. To separate the effects of the prison environment per se from those attributable to a priori dispositions of its inhabitants requires a research strategy in which a "new" prison is constructed, comparable in its fundamental social-psychological milieu to existing prison systems, but entirely populated by individuals who are undifferentiated in all essential dimensions from the rest of society.

Experimental Design

Interpersonal dynamics in a prison environment are to be studied experimentally by designing a functional simulation of a prison in which subjects role-play prisoners and guards for an extended period of time. To assess the power of the social forces on the emergent behavior in this situation, alternative explanations in terms of pre-existing dispositions are to be eliminated through subject selection. A homogeneous "normal" sample is to be chosen after extensive interviewing and diagnostic testing of a large group of volunteer male college students. Half of the subjects are to be randomly assigned to role-play prison guards for eight hours each day, while the others role-play prisoners incarcerated for nearly one full week. Neither group will receive any specific training in these roles. The primary investigator will role-play the prison warden, and consultants from the real prison population (both prisoners and prison officials) will be recruited to assist in the planning and implementation of the prison environment.

Continuous, direct observation of behavioral interactions will be supplemented by video-taped recording, questionnaires, self-report scales and interviews. All these data sources are likely to converge on the conclusion that this simulated prison will develop into a psychologically compelling prison environment.

Significance and Contributions

The authors believe that this demonstration will reveal new dimensions in the social psychology of imprisonment worth pursuing in future research. In addition, this research will provide a paradigm and information base for studying alternatives to existing guard training, as well as for questioning the basic operating principles of which penal institutions rest. There is great need today for prison reforms that recognize the dignity and humanity of both prisoners and guards who are constantly forced into one of the most intimate and potentially deadly encounters known to man. This study has the potential

to inform those reforms.

Works Cited

Small World

"Given any two people In the world , person X and person Z, how many intermediate acquaintance links are needed before 1 and 2 are connected?"

Background

Rationale

Experimental Design

Significance and Contributions

Works Cited

Travers and Milgram 1969 Bernard and Killworth 1978

5.4 *Game-Master Interventions*

Possible Mod: DSMIV compromise

If the game master chooses, he or she may introduce or encourage the development of a compromise position like that found in the DSM-IV:

... although this manual provides a classification of mental disorders, it must be admitted that no definition adequately specifies precise boundaries for the concept of 'mental disorder.' The concept of mental disorder, like many other concepts in medicine and science, lacks a consistent operational definition that covers all situations. All medical conditions are defined on various levels of abstraction—for example, structural pathology (e.g., ulcerative colitis), symptom presentation (e.g., migraine), deviance from a physiological norm (e.g., hypertension), and etiology (e.g., pneumococcal pneumonia). Mental disorders have also been defined by a variety of concepts (e.g., distress, dyscontrol, disadvantage, disability, inflexibility, irrationality, syndromal pattern, etiology, and statistical deviation). Each is a useful indicator for a mental disorder, but none is equivalent to the concept, and different situations call for different definitions."

With something like the 7 part definition included therein:

Features

- A. a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual
- B. is associated with present distress (e.g., a painful symptom) or disability (i.e., impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom

- C. must not be merely an expectable and culturally sanctioned response to a particular event, for example, the death of a loved one
- D. a manifestation of a behavioral, psychological, or biological dysfunction in the individual
- E. neither deviant behavior (e.g., political, religious, or sexual) nor conflicts that are primarily between the individual and society are mental disorders unless the deviance or conflict is a symptom of a dysfunction in the individual

Other Considerations

F no definition adequately specifies precise boundaries for the concept of "mental disorder"

G the concept of mental disorder (like many other concepts in medicine and science) lacks a consistent operational definition that covers all situations

Poor Presidential tone

Not grasping importance of DSMIII

In my experience, and perhaps it is because I'm a Philosopher teaching Philosophy students, students tend towards the radical positions of Albee and Szasz in a way no real, practicing psychologist or psychiatrist would. If this occurs, the game master ought to intervene to help students understand that with the advent of health insurance, if it can't be labeled, you probably won't get paid to treat it.

Failed Research cycle

The game master must ensure that there is at least one research project approved per week, and that that research project has something to report in the week following. If the students on the research committee don't approve any, or there are no proposals from students, I've included a number of examples drawn from historical examples and APA handbooks that can be pulled out and included in the research committee agenda in the appendix Instructors Manual Appendices chapter on p. GM113

Innateness

Chomsky and Piaget actually did debate innateness in 1975 (Piatelli-Palmarini 1984). I included this to point out the fundamental struggle between these different schools of thought. But I'm a Philosopher, and innateness is critical to understanding the distinctions between the various philosophical theories of mind since Plato.

If your game is more focused on the experimental tradition—or the social responsibility of scientists—Piaget can be left out. In that case, Chomsky's written assignment will be on the social responsibilities of social scientists. The role sheet for Chomsky includes this option.

Possible Modification: Voting

I've left the issue of membership in the APA, and the right to vote on many of the main issues, intentionally vague. This is to allow for some flexibility for the instructor. The issue of who can practice mental health treatment is a game issue, realized in competing proposals in 1973. As a corollary then, the issue of membership in the APA may be brought forward. This can also play out as one of the causes of the APA's 'fission' that may occur starting in 1975.

Kameny, Gittings and Gold are not members of the APA at the beginning of the game. If the gamemaster wants everyone to vote, he or she should make clear from the beginning that they are considered to be members of the APA with full rights to vote.

A character does *NOT* need to be a member of the APA to earn credibility. He or she *MUST* be a member of the APA to 'spend' that credibility in the final vote. It follows that non-members must entrust their credibility with a member of the APA so that it can be counted.

6 Debriefing

6.1 What really happened

Actual presidents (Psychological association):

Year	President
1969	George A. Miller, PhD
1970	George W. Albee, PhD
1971	Kenneth B. Clark, PhD—1st African-American, only president born in the Panama Canal zone
1972	Anne Anastasi, PhD—1st woman since 1921
1973	Leona E. Tyler, PhD—Received PhD at youngest age (21) of any APA President
1974	Albert Bandura, PhD—3rd Canadian
1975	Donald T. Campbell, PhD

Table 6.1: Real APA (psychological) Presidents

Hogan, John 1994

Actual history

Little Albert

Contrary to most textbooks that repeat the myth that 'Little Albert' was successfully deconditioned by Watson and was adopted by a family in North Baltimore and went on to live a long, happy, normal life, there is recent evidence to suggest that Little Albert died as a child.

Using photographs from the family and FBI forensics experts, Beck et al. (2009) argue that 'Little Albert' was likely Douglas Merritte, who died at age 6 of acquired hydrocephalus after his mother had left Hopkins. Watson never deconditioned Little Albert, and we simply have no evidence to determine if his conditioned fears remained until his death. Beck, Levinson, and Psychologist

The authors conclude:

None of the folktales we encountered during our inquiry had a factual basis. There is no evidence that the baby's mother was 'outraged,' at her son's treatment or that Douglas's phobia proved resistant to extinction. Douglas was never deconditioned, and he was not adopted by a family north of Baltimore.

Nor was he ever an old man. Our search of seven years was longer than the little boy's life. I laid flowers on the grave of my longtime 'companion' turned, and simultaneously felt a great peace and profound loneliness. Beck, Levinson, and Psychologist

The APA's Monitor carried the story January 2010 (DeAngelis 2010)

There were a number of blog posts that followed, including: <http://mindhacks.com/2009/10/22/little-albert-lost-and-found>. And a BBC special: <https://www.youtube.com/watch?v=KJnJ1Q8PAJk>

In 2012, however, a study by Nancy Digdon and Russell Powell called Beck's findings into doubt. He found that there was another child at Hopkins at the time named William Albert Barger who fits the description of Little Albert—and he went on to have a normal adult life.¹ Powell, Digdon, and Psychologist

¹ See <https://www.chronicle.com/interactives/littlealbert>.

Socarides and Bieber

The dates I used in the introduction are not exact, although I've tried to remain faithful where I could. Part of the problem here is that many of the stories that are circulated, even by eyewitnesses, vary. Ron Gold, for example, dates the events in Hawaii to 1972, but the APA records the 1972 conference in Dallas, and the Hawaii conference in 1973.

The Dr. H. Anonymous speech actually happened in 1972 during a symposium titled "Psychiatry: Friend or Foe to Homosexuals: A Dialogue." It also featured Frank Kameny, Barbara Gittings and Judd Marmor in addition to John Fryer.

In 1971, Kameny, and Gittings held a panel titled "Lifestyles of non-patient homosexuals," which the participants joking referred to as "Lifestyles of impatient homosexuals." It featured Larry Littlejohn of the Society for Individual Rights in San Francisco, Del Martin, a founder the Lesbian activist organization Daughters of Bilitis, Lilli Vincenz,² another lesbian activists, and Jack Baker, the gay president-elect of the student body of the University of Minnesota. I've found a number of websites and personal stories that get these two distinct events confused. At the risk of perpetuating these confusions, I've combined them into a single event in 1971.

² There are a number of photos of both Del Martin and Lilli Vincenz in the Barbara Gittings and Kay Tobin Lahusen gay history papers and photographs archive of the NY Public Library (<https://digitalcollections.nypl.org/collections/barbara-gittings-and-kay-tobin-lahusen-gay-history-papers?tab=navigation>)



Figure 6.1: Manuscripts and Archives Division, The New York Public Library. (1972 - 1972). *Gittings, Kameny, Dr. John E. Fryer, Dr. Judd Marmor on panel*
Retrieved from [urlhttp://digitalcollections.nypl.org/items/510d47e3-830b-a3d9-e040-e00a18064a99](http://digitalcollections.nypl.org/items/510d47e3-830b-a3d9-e040-e00a18064a99)

Socarides petitioned the New York District branch of the APA for his task force on sexual deviation, not the national organization as in our game. His petition was granted and his report filed in 1972. The leadership then rejected his report, on the grounds that its basis in psychoanalytic theory was unacceptable.

Complicating matters was the existence of NARTH—the National Association for Research and Treatment of Homosexuality—which has been adopted by many on the religious right in recent years. In 1995, Socarides gave an interview to the NY Times, promoting his book *Homosexuality: A Freedom Too Far - A Psychoanalyst Answers 1,000 Questions About Causes and Cure and the Impact of the Gay Rights Movement on American Society*, in which he claimed that the declassification occurred because of political pressure and not scientific evidence. The themes of the interview will be familiar to anyone with even tangential knowledge of the rhetoric of the far right: persecution of conservatives, liberal media bias, gay agendas and conspiracies, etc. Charles Socarides' son, Richard Socarides, who is gay, became a gay activist, advised President Clinton on LGBT issues, and ultimately was elected to the presidency of the activist organization 'Equality Matters.' ³ That fact has lead many defenders of Socarides to accuse any press figure who mentions it of making an 'ad hominem' attacks against Socarides.

³ He is also an alumnus of my undergraduate college, Antioch.

NARTH continued until 2017 under the directorship of Socarides' student Joseph Nicolosi. Nicolosi passed away on March 8th, 2017, and NARTH shut down as a consequence.

Dr. Joseph Nicolosi, Sr. (1947 - 2017)

We are deeply saddened to announce that Dr. Joseph Nicolosi, Sr. passed away unexpectedly on March 8, 2017.

The Thomas Aquinas Psychological Clinic has therefore closed down its psychotherapy practice.

Although our practice is closed, there are several other therapists in their own private practices, trained by Dr. Nicolosi, to whom we can refer you. Please email our office for information. These therapists share your worldview and support your sense of who you were meant to be. You can email us at tapc1@earthlink.net.

We will also continue to make josephnicolosi.com available to anyone interested in understanding Dr. Nicolosi's lifetime of work and thought. The articles he wrote on this web site have timeless value. He is the originator of reparative therapy, and these articles, as well as the four books he wrote, will live on in the history of psychotherapy.

To read tributes to Dr. Nicolosi's life and work from friends and colleagues, please click on this link: <http://www.forevermissed.com/dr-joseph-nicolosi-sr/#about>. If you knew Dr. Nicolosi or were familiar with his work, you may add your own tribute there.

Sincerely,
The Office Staff

The UK Newspaper 'The Independent' ran a story in January 2010 by Patrick Strudwick about his experience in gay-to-straight conversion program inspired by Nicolosi. The story created a national concern about psychiatrists in the UK who still may be offering psychotherapy to 'treat' or 'cure' homosexuality. The story is available here: <http://www.independent.co.uk/life-style/health-and-families/features/the-exgay-files-the-bizarre-world-of-gaytostraight-conversion-1884947.html>

Figure 6.2: Screenshot of message posted on NARTH homepage after death of Nicolosi

Leona Tyler Principle /Goldwater Rule

Following the Goldwater affair, the American Psychiatric Association added section 7.3 to their "The Principles of Medical Ethics With Annotations Especially Applicable to Psychiatry," which states:

On occasion psychiatrists are asked for an opinion about an individual who is in the light of public attention or who has disclosed information about himself/herself through public media. In such circumstances, a

psychiatrist may share with the public his or her expertise about psychiatric issues in general. However, it is unethical for a psychiatrist to offer a professional opinion unless he or she has conducted an examination and has been granted proper authorization for such a statement.

This section has become known as 'The Goldwater Rule.' For a full history, see Mayer, John D. (2010) "The Goldwater Rule: The rationale of the Goldwater Rule" Psychology Today Blog (<http://www.psychologytoday.com/blog/the-personality-analyst/201005/the-goldwater-rule>)

The American Psychological Association adopted 'Leona Tyler principle' was adopted in 1973 by the American Psychological Association, and still holds today. It states:

As citizens, members of the APA have the right to advocate for any cause through the myriad of political advocacy organizations, but when psychologists and psychiatrists speak for the profession through APA public stances and proclamations, it should be from science and professional experience.

The principle was named for Leona Tyler simply because she was president at the time, not necessarily because she advocated for it, as she does in this game.

There was a flurry of discussion about the Goldwater rule and the Leona Tyler Principle during and after the 2016 Presidential Election of Donald Trump. Notably, the American Psychoanalysts Association lifted its equivalent rule so the membership could speak about Trump's mental health (<https://gizmodo.comamerican-psychanalytic-association-says-its-members-ar-1797247879>)

Demedicalization

The story of the demedicalization of homosexuality as a mental illness has been retold a couple of times in recent years. Famously, Alix Spiegel, grandson of the actual John P. Spiegel, produced a version of the story for NPR's show 'This American Life'. The episode, named '81 words' is available online at <http://www.thisamericanlife.org/radio-archives/episode/204/81-words>. It is an excellent retelling of these events, and partially inspired this game. Ron Gold's retelling of the event is available here: <http://www.queerstories.org/custom.html>.

In the role sheet for Spitzer, I express some doubts about the standard story told by Spiegel, because Spitzer's timing and motivation for inserting himself into the controversy doesn't seem to match up for a normal person. However, as I was to learn, Spitzer is not a 'normal' person.

In a 2005 issue of The New Yorker Magazine, Spiegel writes that:

Despite Spitzer's genius at describing the particulars of emotional behavior, he didn't seem to grasp other people very well. Jean Endicott,

his collaborator of many years, says, "He got very involved with issues, with ideas, and with questions. At times he was unaware of how people were responding to him or to the issue. He was surprised when he learned that someone was annoyed. He'd say, 'Why was he annoyed? What'd I do?'" After years of confrontations, Spitzer is now aware of this shortcoming, and says that he struggles with it in his everyday life. "I find it very hard to give presents," he says. "I never know what to give. A lot of people, they can see something and say, 'Oh, that person would like that.' But that just doesn't happen to me. It's not that I'm stingy. I'm just not able to project what they would like." Frances argues that Spitzer's emotional myopia has benefitted him in his chosen career: "He doesn't understand people's emotions. He knows he doesn't. But that's actually helpful in labeling symptoms. It provides less noise."

If this is correct, I am wrong in my assertion that it was implausible for the Chair of the Nomenclature Committee to attend a speech in New York by Socarides' without expecting a demonstration. If Spitzer, by his own account, can't imagine what someone else would like as a gift, how can we expect him to imagine how much anger there was in the Gay community towards Socarides?

After much thought, I decided to leave my assertion in the role sheet, because this is 'Reacting' to the past, not "Recreating" the past. A student playing the part needs to think himself or herself into Spitzer's *position*, not his *mind*—and whatever unique mind Robert Spitzer has, it is probably uncommon in undergraduates.

Marr

David Marr joins Miller and Chomsky at MIT in 1977 and died in 1980 of Leukemia. His posthumously published book *Vision* is still standard reading in Cognitive Science courses.

Spitzer Task force

See ch 3 of Kutchins and Kirk 1997.

Homodysphilia was proposed by Spitzer without consultation with the subcommittee on human sexuality. Green sent Spitzer a strongly worded criticism of both the classification and the method Spitzer used to introduce it and resigns in protest. Spitzer constructs a 'survey' instrument that he sends out to the entire taskforce and finds that there isn't enough agreement on the issue to leave it with the subcommittee, and presents the idea to the entire taskforce, thus outmaneuvering Green. The committee narrowly approves 'ego-dystonic homosexuality', but isn't thrilled. Judd Marmor, who voted in favor, writes later that he would have preferred no mention of homosexuality, but politics required it.

Aversion Therapy

(from wikipedia) Since 1994, the American Psychological Association has declared that aversion therapy is a dangerous practice that does not work. Since 2006, the use of aversion therapy to treat homosexuality has been in violation of the codes of conduct and professional guidelines of the American Psychological Association and American Psychiatric Association. The use of aversion therapy to treat homosexuality is illegal in some countries. The standard in psychotherapy in America and Europe is currently Gay Affirmative Psychotherapy. Guidelines for Gay Affirmative Psychotherapy can be found by APA.

From Issues in Psychotherapy with Lesbian and Gay Men: A Survey of Psychologists. Linda Garnets, Los Angeles, CA Kristin A. Hancock, Berkeley, CA Susan D. Cochran, California State University, Northridge Jacqueline Goodchilds, University of California, Los Angeles Letitia Anne Peplau, University of California, Los Angeles. Report of the APA. <http://search.apa.org/search?facet=classification%3aSexuality%20classification%3aTherapy&query=Aversion%20Therapy>

In 1975, the American Psychological Association (APA) took a strong stance regarding bias toward lesbians and gay men, resolving that homosexuality per se implies no impairment in judgment, reliability or general social and vocational abilities (see Appendix A for the full text of the resolution). The APA urged psychologists to take the lead in removing the stigma of mental illness long associated with homosexual orientations (Conger, 1975). In recent years, attention has been drawn to ways in which a client's ethnicity, gender, sexual orientation, or physical disability can affect clinical judgment and treatment strategies. There has been a corresponding effort to develop guidelines to help practitioners avoid bias in psychotherapy (APA, 1975). Recognizing that practice does not spontaneously or quickly follow policy changes, the Committee on Lesbian and Gay Concerns (CLGC), sponsored jointly by the Board of Social and Ethical Responsibility in Psychology (BSERP) and the Board of Professional Affairs (BPA), formed a task force in 1984 to investigate the range of bias that may occur in psychotherapy with lesbians and gay men. This article is an abridged report of the task force's research, findings, and recommendations.

Reparative therapy

See discussion of NARTH in Demedicalization subsubsection 31.3 on p. R177 above.

The American Psychological Association rejected reparative therapy most recently in 2009. The American Psychiatric Association in 2000, and the American Psychoanalytic Association in 2012. For the full list, see the Human Rights Campaign's list: <https://www.hrc.org/resources/policy-and-position-statements-on-conversion-therapy>

Evelyn Hooker

E. Hooker was given the 1991 Award for Distinguished Contribution to Psychology in the Public Interest, presented by the American Psychological Association. The citation read:

"When homosexuals were considered to be mentally ill, were forced out of government jobs, and were arrested in police raids, Evelyn Hooker courageously sought and obtained research support from the National Institute of Mental Health (NIMH) to compare a matched sample of homosexual and heterosexual men. Her pioneering study, published in 1957, challenged the widespread belief that homosexuality is a pathology by demonstrating that experienced clinicians using psychological tests widely believed at the time to be appropriate could not identify the nonclinical homosexual group. This revolutionary study provided empirical evidence that normal homosexuals existed, and supported the radical idea then emerging that homosexuality is within the normal range of human behavior. Despite the stigma associated with homosexuality, she received an NIMH Research Career Award in 1961 to continue her work. In 1967, she became chair of the NIMH Task Force on Homosexuality, which provided a stamp of validation and research support for other major empirical studies. Her research, leadership, mentorship, and tireless advocacy for an accurate scientific view of homosexuality for more than three decades has been an outstanding contribution to psychology in the public interest."

She reflects on the award here: http://psychology.ucdavis.edu/rainbow/html/hooker_address.html I relied heavily on this reflection in the writing the role sheet

1997 resolution: <http://psychology.ucdavis.edu/rainbow/HTML/resolution97.html>

6.2 After this era

Margaret Mead debunked

Are Ebbinghaus' experiments replicable?

Jaap M J Murre and Joeri Dros. 2015. "Replication and Analysis of Ebbinghaus' Forgetting Curve". *PLOS ONE* 10, no. 7 (): e0120644

Abortion

"APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676a–676

"APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676c–677

"APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676b–676

- “Position Statement on Abortion”. 1979. *American Journal of Psychiatry* 136, no. 2 (): 272–272
- “Position Statement on Abortion”. 1970. *American Journal of Psychiatry* 126, no. 10 (): 1554–1554
- “Position Statement on Abortion”. 1967. *American Journal of Psychiatry* 124, no. 3 (): 450-a–450

The DSM-V Controversy

These were collected in 2009. The issue is ongoing.

NIMH Rejection: <http://www.newyorker.com/online/blogs/elements/2013/05/the-new-criteria-for-mental-disorders.html>

The D.S.M. And the Nature of Disease <http://www.newyorker.com/online/blogs/elements/2013/04/the-dsm-and-the-nature-of-disease.html>

(2009). “What Should Count as a Mental Disorder in DSM-V? - Psychiatric Times.” from <http://www.psychiatrictimes.com/display/article/10168/1402032>.

Andrews, G. (2009). Stress-induced and fear circuitry disorders : advancing the research agenda for DSM-V. Washington, DC, American Psychiatric Pub.

Beach, S. R. H. (2006). Relational processes and DSM-V : neuroscience, assessment, prevention, and intervention. Washington, DC, American Psychiatric Pub.

Kal, E. F. (2008). “DSM-V Compromise Proposed.” *Psychiatr News* 43(21): 23-a-23-a.

O'Donohue, W. T., K. A. Fowler, et al. (2007). Personality disorders : toward the DSM-V. Los Angeles, SAGE Publications.

Porcelli, P. and N. Sonino (2007). Psychological factors affecting medical conditions : a new classification for DSM-V. Basel ; New York, Karger.

Saunders, J. B. and American Psychiatric Association. (2007). Diagnostic issues in substance use disorders : refining the research agenda for DSM-V. Arlington, VA, American Psychiatric Association.

Spitzer, R. L. (2008). “DSM-V: Open and Transparent?” *Psychiatr News* 43(14): 26–26.

Sunderland, T. and American Psychiatric Association. (2007). Diagnostic issues in dementia : advancing the research agenda for DSM-V. Arlington, VA, American Psychiatric Association.

Widiger, T. A. (2006). Dimensional models of personality disorders : refining the research agenda for DSM-V. Washington, D.C., American Psychiatric Association.

Spitzer 2003 Controversy

In 2003, Spitzer said that it is possible to change one's sexual orientation:

Spitzer, R. (2003). "Can some gay men and lesbians change their sexual orientation? 200 participants reporting a change from homosexual to heterosexual orientation." *Archives of Sexual Behavior* 32(5): 403–417.

It caused an uproar. See:

Epstein, R. (2003). "Am I Anti-Gay? You be the judge. A letter from the editor in chief." *Psychology Today* Jan/Feb.

Hausman, K. (2001). "Furor Erupts Over Study On Sexual Orientation." *Psychiatr News* 36(13): 20–34.

Nicolosi, J. (2007, 4/30/2007). "Why I Am Not a Neutral Therapist." Retrieved 10/24/2008, from <http://www.narth.com/docs/notneutral.html>.

Nicolosi, J. (2008). "The Meaning of Same-Sex Attraction." 10/24/2008, from <http://www.narth.com/docs/niconew.html>.

Satinover, J. (?). "The "Trojan Couch": How the Mental Health Associations Misrepresent Science." National Association for Research and Therapy of Homosexuality. Retrieved 10/24, 2008, from <http://www.narth.com/docs/TheTrojanCouchSatinover.pdf>.

Zucker, K. and J. Drescher, Eds. (2006). *Ex-Gay Research: Analyzing the Spitzer Study and Its Relation to Science, Religion, Politics, and Culture*. New York, Routledge.

7 Works Cited

References

- "APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676a–676.
- "APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676c–677.
- "APA position statement on abortion". 1993. *American Journal of Psychiatry* 150, no. 4 (): 676b–676.
- Beck, H P, S Levinson, and G Irons American Psychologist. "Finding little Albert: A journey to John B. Watson's infant laboratory." psycnet.apa.org.
- Bernard, H Russell, and Peter D Killworth. 1978. "A Review of the Small World Literature." () .
- Cacioppo, John, and Laura Freberg. 2018. *Discovering Psychology: The Science of Mind*. Cengage Learning.
- Cervone, D, and T L Caldwell. 2015. *Psychology: The science of person, mind, and brain*.
- Clabough, Erin B D, and Seth W Clabough. 2016. "Using Rubrics as a Scientific Writing Instructional Method in Early Stage Undergraduate Neuroscience Study". *The Journal of Undergraduate Neuroscience Education* 15, no. 1 (): A85–A93.
- Crissey, Marie Skodak, and George W Albee. 1982. "The 1982 division 27 award for distinguished contributions to community psychology and community mental health: George W. Albee". *American Journal of Community Psychology* 10, no. 1 (): 1–36.
- DeAngelis, T. 2010. 'Little Albert' regains his identity.
- Gatz, Margaret. 2007. "In memory of George Albee". *The Journal of Primary Prevention* 28, no. 1 (): 19–20.
- Gazzaniga, Michael. 2018. *Psychological Science*. Sixth Edition. W. W. Norton & Company.

- Goode, Erica. 2001. "Anne Anastasi, the 'Test Guru' of Psychology, Is Dead at 92". *The New York Times* ().
- Haney, Craig, Curtis Banks, and Philip Zimbardo. 1973. "Interpersonal Dynamics in a Simulated Prison". *International Journal of Criminology and Penology* 1:66–97.
- Helmore, Edward. 1996. "Obituary: Kurt Freund". *The Independent* ().
- Hockenbury, Sandra E, Susan A Nolan, and Don H Hockenbury. 2014. *Psychology*. 7th. Worth Publishers.
- Hogan, John. 1994. "Thousands Apply to U.S. to Forgive Their Student Loans, Saying Schools Defrauded Them". *Ann N Y Acad Sci*. 30 (727): 133–138.
- Hooker, Evelyn. 1956. "A Preliminary Analysis of Group Behavior of Homosexuals". *The Journal of Psychology* 42, no. 2 (): 217–225.
- Hothersall, David. 1984. *History of Psychology*. New York: Random House.
- Hummer, Andy. 2017. *Ron Gold, Pioneer in Challenging Sickness Label, Dies*. <https://www.gaycitynews.nyc/stories/2017/10/w27290-ron-gold-pioneer-challenging-sickness-label-dies-2017-05-16.html>.
- Kutchins, Herb, and Stuart Kirk. 1997. *Making us crazy : DSM : the psychiatric bible and the creation of mental disorders*.
- Lilienfeld, Scott O, et al. 2014. *Psychology*. 3rd. From Inquiry to Understanding. Prentice Hall.
- Milgram, Stanley. 1967. *The small-world problem*.
- Murre, Jaap M J, and Joeri Dros. 2015. "Replication and Analysis of Ebbinghaus' Forgetting Curve". *PLOS ONE* 10, no. 7 (): e0120644.
- Myers, David G, and Charles K DeWall. 2015. *Psychology*. Worth Publishers.
- Piatelli-Palmarini. 1984. "Play, Dreams, and Imitation in Childhood".
- "Position Statement on Abortion". 1967. *American Journal of Psychiatry* 124, no. 3 (): 450-a–450.
- "Position Statement on Abortion". 1970. *American Journal of Psychiatry* 126, no. 10 (): 1554–1554.
- "Position Statement on Abortion". 1979. *American Journal of Psychiatry* 136, no. 2 (): 272–272.
- Powell, R A, N Digdon, and B Harris American Psychologist. "Correcting the record on Watson, Rayner, and Little Albert: Albert Barger as "Psychology's lost boy""." psycnet.apa.org.

- Price, Richard H. 2007. "George Albee: A man with a better idea". *The Journal of Primary Prevention* 28, no. 1 (): 55–57.
- Rathus, Spencer. 2012. *Psychology: Concepts & Connections, Brief Version*. Cengage Learning.
- Rosenthal, David L. 1973. "On Being Sane In Insane Places". *Science* 179:250–258.
- Shepard, Roger N, and Jacqueline Metzler. 1971. "Mental Rotation of Three-Dimensional Objects". *Science* 171 (3972): 701–703.
- Siegelman, Marvin. 1972. "Adjustment of male homosexuals and heterosexuals". *Archives of Sexual Behavior* 2 (1): 9–25.
- Spitzer, Robert, and J L Fleiss. 1974. "A Re-analysis of the Reliability of Psychiatric Diagnosis". *The British Journal of Psychiatry* 125, no. 4 (): 341–347.
- Travers, Jeffrey, and Stanley Milgram. 1969. "An Experimental Study of the Small World Problem". *Sociometry* 32, no. 4 (): 425–443.
- Zimbardo, P G, R Johnson, and V McCann. 2012. *Psychology: Core Concepts*. 7th. Pearson.

8 Indicies

Instructor's Manual Index

- Albee, George, GM17
Anastasi, Anne, GM19
- Bandura, Albert, GM59
Bieber, Irving, GM29
- Chomsky, Noam, GM55
Clark, Kenneth, GM19
Clark, Kenneth: Proposal,
GM34
- DSM-I, GM11
DSM-II, GM68
DSM-III, GM13
- Fordney-Settlage, Diane,
GM47
Freund, Kurt, GM19
Fryer, John, GM96
- Gebhard, Paul, GM73
Gittings, Barbara, GM22
Gold, Ron, GM96
Green, Richard, GM74
- Harlow, Harry, GM19
Hooker, Evelyn, GM72
Hopcke, Robert, GM29
- Kameny, Frank, GM39
- Leif, Harold, GM19
- Marmor, Judd, GM19
Marr, David, GM19
Milgram, Stanley, GM19
Milgrim, Stanley, GM91
Miller, George, GM38
- Piaget, Jean, GM92
- Socarides, Charles, GM12
Spiegel, John, GM39
Spitzer Taskforce, GM100
Spitzer, Robert, GM13
Spitzer, Robert: 2003
Controversy, GM104
Szasz, Thomas, GM92
- Tyler, Leona, GM19
Tyler, Leona: Leona Tyler
Principle, GM98
- Zimbardo, Philip, GM35
Zimbardo, Philip: Research,
GM46

Full index

Game book page numbers are in plain arabic, Instructor's are preceded by 'GM' and Rolesheets by 'R'.

Instructors Manual Appendices

A *Real-world papers*

The real-world examples are summaries of existing psychological experiments during the era. They were written by my students as a part of gameplay over a couple of instances, and edited and revised by myself.

Mental rotation



Mental Rotation of Three-Dimensional Objects

Author(s): Roger N. Shepard and Jacqueline Metzler

Source: *Science*, New Series, Vol. 171, No. 3972 (Feb. 19, 1971), pp. 701-703

Published by: American Association for the Advancement of Science

Stable URL: <http://www.jstor.org/stable/1731476>

Accessed: 05/04/2010 11:40

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aaas>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

solution is shown for day 1 only (16), because, even though the intake was relatively small (mean intake, 98 ml), the cats became ill. Two drank sucrose almost continuously up to the criterion and subsequently vomited and developed diarrhea. The others did not vomit but developed diarrhea. The illness apparently led to conditioned aversion. After a week's rest, cats rejected 0.375M sucrose (mean intake, 18 ml). This same thing happened with 0.5M sucrose solution and 24-hour intake.

Frings's (9) finding that sucrose in dilute milk (one part milk to four parts water) is preferred by cats fits in well with the result presented here. Mean sodium and chlorine content for whole milk so diluted would approximate 0.006M NaCl (17). The exact whole-mouth salivary NaCl concentration for the cat is not known, but it must fall between 0.01M and 0.16M NaCl (18). For adapting concentrations in this range, electrophysiological data (3) suggest that the 0.006M NaCl in the milk used by Frings would at least partially suppress the water-after-NaCl response.

The taste of water has been widely ignored in behavioral testing. It is now clear that water should be regarded not as a neutral solvent but rather as a taste stimulus itself. The implications for receptor mechanisms are still unclear. Water appears to produce some responses by removing other stimuli (2), but it may also stimulate receptors directly [see (19) for a review of various structural models of water]. Nevertheless, electrophysiological studies can suggest how water tastes can be manipulated to assess the taste of any given substance in water.

L. M. BARTOSHUK

*John B. Pierce Foundation Laboratory,
290 Congress Avenue,
New Haven, Connecticut 06519*

M. A. HARND

L. H. PARKS

*Pioneering Research Laboratory,
U.S. Army Natick Laboratories,
Natick, Massachusetts 01760*

References and Notes

- G. Liljestrand and Y. Zotterman, *Acta Physiol. Scand.* **32**, 291 (1954); Y. Zotterman, *ibid.* **37**, 60 (1956); — and H. Diamant, *Nature* **183**, 191 (1959).
- M. J. Cohen, S. Hagiwara, Y. Zotterman, *Acta Physiol. Scand.* **33**, 316 (1955).
- L. M. Bartoshuk, thesis, Brown University (1965).
- and C. Pfaffmann, *Fed. Proc.* **24**, 207 (1965); C. Pfaffmann, in *Olfaction and Taste* C. Pfaffmann, Ed. (Rockefeller Univ. Press, New York, 1969), vol. 3, p. 528.
- L. M. Bartoshuk, D. H. McBurney, C. Pfaffmann, *Science* **143**, 967 (1964); L. M. Bartoshuk, *Percept. Psychophys.* **3**, 69 (1968); D. H. McBurney, *J. Exp. Psychol.* **72**, 869 (1966); in *Olfaction and Taste*, C. Pfaffmann, Ed. (Rockefeller Univ. Press, New York, 1969), vol. 3, p. 407.
- O. Maller and M. R. Kare, *Anim. Behav.* **15**, 8 (1967); V. G. Dethier and M. V. Rhoades, *J. Exp. Zool.* **126**, 177 (1954); C. Duncan, *Anim. Behav.* **8**, 54 (1960); H. L. Jacobs and M. L. Scott, *Poultry Sci.* **36**, 8 (1957); B. P. Halpern, *Amer. J. Physiol.* **203**, 541 (1962); M. R. Kare and M. S. Ficken, in *Olfaction and Taste*, Y. Zotterman, Ed. (Pergamon, New York, 1963), vol. 2, p. 285; J. Ganchrow and G. L. Fisher, *Psychol. Rep.* **22**, 503 (1968); C. Pfaffmann, *Amer. Psychol.* **20**, 21 (1965).
- J. A. Carpenter, *J. Comp. Physiol. Psychol.* **49**, 139 (1956).
- Y. Zotterman, *Skand. Archiv Physiol.* **72**, 73 (1935); C. Pfaffmann, *J. Cell. Comp. Physiol.* **17**, 243 (1941).
- H. Frings, *Experientia (Basel)* **7**, 424 (1951).
- C. Pfaffmann, *J. Neurophysiol.* **18**, 492 (1955).
- The sensitivities to NaCl, QHCl, sucrose, and HCl appear to be independently associated in accordance with the random distribution hypothesis of M. Frank and C. Pfaffmann [*Science* **164**, 1183 (1969)]. In addition, the four contingencies producing water responses also appear to be independently associated in accordance with this hypothesis.
- The correlation coefficient for water-after-NaCl responses and NaCl responses was -0.83 ($P < .005$). Only those fibers were included for which at least one response met the criterion. Previous reports of a negative correlation between water responses and NaCl responses [Cohen et al. (2); J. Nagaki, S. Yamashita, M. Sato, *Jap. J. Physiol.* **14**, 67 (1964)] probably reflect the negative correlation between water-after-NaCl and NaCl responses since the rinse was Ringer solution. The correlation coefficients for NaCl and water-after-QHCl, water-after-sucrose, and water-after-HCl are -0.65 ($P < .05$), -0.54 ($P > .05$), and -0.28 ($P > .2$), respectively.
- We thank M. Dvorak and O. Stark of the Food Sciences Laboratory at the Natick Army Laboratories for analyses of the atomic absorption spectra of tap and distilled water samples. Samples were analyzed for Na, K, Ca, Mg, and halide. Tap water checked weekly for 2 months remained relatively constant with mean levels of $9.6 \times 10^{-4}M$ Na, $6.8 \times 10^{-5}M$ K, $7.1 \times 10^{-4}M$ Ca, $2.2 \times 10^{-4}M$ Mg, and $1.0 \times 10^{-3}M$ halide (predominantly Cl). Distilled water contained $8.7 \times 10^{-7}M$ Na, $5.1 \times 10^{-6}M$ K, $2.5 \times 10^{-6}M$ Ca, and $4.1 \times 10^{-7}M$ Mg.
- P. Hore and M. Messer, *Comp. Biochem. Physiol.* **24**, 717 (1968).
- S. Siegel, *Nonparametric Statistics* (McGraw-Hill, New York, 1956).
- Since testing was discontinued after 1 day, the position of sucrose was not counterbalanced. To ensure that position preference could not account for the results at a sucrose concentration of 0.375M, position preference under all other conditions was tested with a two-tailed Walsh test (15). The results were not significant ($P > .1$).
- K. Diem, Ed., *Scientific Tables* (Geigy Pharmaceuticals, Ardsley, N.Y., ed. 6, 1962), p. 516.
- L. H. Schneyer and C. A. Schneyer, in *Handbook of Physiology: Alimentary Canal*, C. F. Code, Ed. (American Physiological Society, Washington, D.C., 1967), vol. 2.
- O. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford Univ. Press, New York, 1969).
- Electrophysiological data and preliminary behavioral data were obtained at Brown University where work was supported in part by a PHS predoctoral fellowship to L.M.B. and NSF grants G-14332 (to C. Pfaffmann) and GB-2754 (to C. Pfaffmann and L.M.B.). Final behavioral data were collected at the U.S. Army Natick Laboratories. We thank R. L. Gentile, J. C. Stevens, L. E. Marks, and W. S. Cain for valuable comments on the manuscript.

19 October 1970; revised 13 November 1970 ■

Mental Rotation of Three-Dimensional Objects

Abstract. The time required to recognize that two perspective drawings portray objects of the same three-dimensional shape is found to be (i) a linearly increasing function of the angular difference in the portrayed orientations of the two objects and (ii) no shorter for differences corresponding simply to a rigid rotation of one of the two-dimensional drawings in its own picture plane than for differences corresponding to a rotation of the three-dimensional object in depth.

Human subjects are often able to determine that two two-dimensional pictures portray objects of the same three-dimensional shape even though the objects are depicted in very different orientations. The experiment reported here was designed to measure the time that subjects require to determine such identity of shape as a function of the angular difference in the portrayed orientations of the two three-dimensional objects.

This angular difference was produced either by a rigid rotation of one of two identical pictures in its own picture plane or by a much more complex, nonrigid transformation, of one of the pictures, that corresponds to a (rigid) rotation of the three-dimensional object in depth.

This reaction time is found (i) to

increase linearly with the angular difference in portrayed orientation and (ii) to be no longer for a rotation in depth than for a rotation merely in the picture plane. These findings appear to place rather severe constraints on possible explanations of how subjects go about determining identity of shape of differently oriented objects. They are, however, consistent with an explanation suggested by the subjects themselves. Although introspective reports must be interpreted with caution, all subjects claimed (i) that to make the required comparison they first had to imagine one object as rotated into the same orientation as the other and that they could carry out this "mental rotation" at no greater than a certain limiting rate; and (ii) that, since they perceived the two-dimensional pictures as objects

in three-dimensional space, they could imagine the rotation around whichever axis was required with equal ease.

In the experiment each of eight adult subjects was presented with 1600 pairs of perspective line drawings. For each pair the subject was asked to pull a right-hand lever as soon as he determined that the two drawings depicted objects of different three-dimensional shapes. According to a random sequence, in half of the pairs (the "same" pairs) the two objects could be rotated into congruence with each other (as in Fig. 1, A and B), and in the other half (the "different" pairs) the two objects differed by a reflection as well as a rotation and could not be rotated into congruence (as in Fig. 1C).

The choice of objects that were mirror images or "isomers" of each other for the "different" pairs was intended to prevent subjects from discovering

some distinctive feature possessed by only one of the two objects and thereby reaching a decision of noncongruence without actually having to carry out any mental rotation. As a further precaution, the ten different three-dimensional objects depicted in the various perspective drawings were chosen to be relatively unfamiliar and meaningless in overall three-dimensional shape.

Each object consisted of ten solid cubes attached face-to-face to form a rigid armlike structure with exactly three right-angled "elbows" (see Fig. 1). The set of all ten shapes included two subsets of five: within either subset, no shape could be transformed into itself or any other by any reflection or rotation (short of 360°). However, each shape in either subset was the mirror image of one shape in the other subset, as required for the construction of the "different" pairs.

For each of the ten objects, 18 different perspective projections—corresponding to one complete turn around the vertical axis by 20° steps—were generated by digital computer and associated graphical output (1). Seven of the 18 perspective views of each object were then selected so as (i) to avoid any views in which some part of the object was wholly occluded by another part and yet (ii) to permit the construction of two pairs that differed in orientation by each possible angle, in 20° steps, from 0° to 180°. These 70 line drawings were then reproduced by photo-offset process and were attached to cards in pairs for presentation to the subjects.

Half of the "same" pairs (the "depth" pairs) represented two objects that differed by some multiple of a 20° rotation about a vertical axis (Fig. 1B). For each of these pairs, copies of two appropriately different perspective views were simply attached to the cards in the orientation in which they were originally generated. The other half of the "same" pairs (the "picture-plane" pairs) represented two objects that differed by some multiple of a 20° rotation in the plane of the drawings themselves (Fig. 1A). For each of these, one of the seven perspective views was selected for each object and two copies of this picture were attached to the card in appropriately different orientations. Altogether, the 1600 pairs presented to each subject included 800 "same" pairs, which consisted of 400 unique pairs (20 "depth" and 20 "picture-plane" pairs at each of the ten angular differences from 0° to 180°), each of which was

presented twice. The remaining 800 pairs, randomly intermixed with these, consisted of 400 unique "different" pairs, each of which (again) was presented twice. Each of these "different" pairs corresponded to one "same" pair (of either the "depth" or "picture-plane" variety) in which, however, one of the three-dimensional objects had been reflected about some plane in three-dimensional space. Thus the two objects in each "different" pair differed, in general, by both a reflection and a rotation.

The 1600 pairs were grouped into blocks of not more than 200 and presented over eight to ten 1-hour sessions (depending upon the subject). Also, although it is only of incidental interest here, each such block of presentations was either "pure," in that all pairs involved rotations of the same type ("depth" or "picture-plane"), or "mixed," in that the two types of rota-

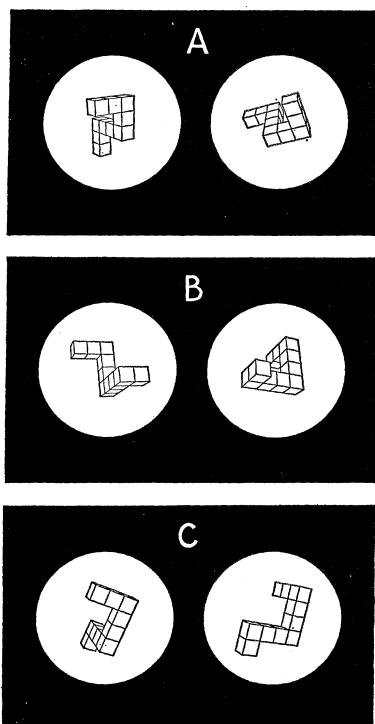


Fig. 1. Examples of pairs of perspective line drawings presented to the subjects. (A) A "same" pair, which differs by an 80° rotation in the picture plane; (B) a "same" pair, which differs by an 80° rotation in depth; and (C) a "different" pair, which cannot be brought into congruence by any rotation.

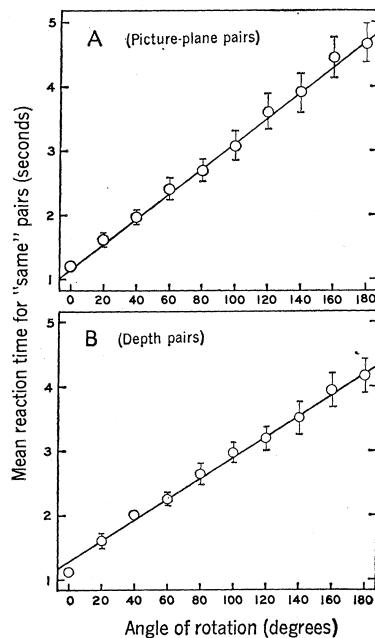


Fig. 2. Mean reaction times to two perspective line drawings portraying objects of the same three-dimensional shape. Times are plotted as a function of angular difference in portrayed orientation: (A) for pairs differing by a rotation in the picture plane only; and (B) for pairs differing by a rotation in depth. (The centers of the circles indicate the means and, when they extend far enough to show outside these circles, the vertical bars around each circle indicate a conservative estimate of the standard error of that mean based on the distribution of the eight component means contributed by the individual subjects.)

tion were randomly intermixed within the same block.

Each trial began with a warning tone, which was followed half a second later by the presentation of a stimulus pair and the simultaneous onset of a timer. The lever-pulling response stopped the timer, recorded the subject's reaction time and terminated the visual display. The line drawings, which averaged between 4 and 5 cm in maximum linear extent, appeared at a viewing distance of about 60 cm. They were positioned, with a center-to-center spacing that subtended a visual angle of 9°, in two circular apertures in a vertical black surface (see Fig. 1, A to C).

The subjects were instructed to respond as quickly as possible while keeping errors to a minimum. On the average only 3.2 percent of the responses were incorrect (ranging from 0.6 to 5.7 percent for individual subjects). The reaction-time data presented below include only the 96.8 percent correct responses. However, the data for the incorrect responses exhibit a similar pattern.

In Fig. 2, the overall means of the reaction times as a function of angular difference in orientation for all correct (right-hand) responses to "same" pairs are plotted separately for the pairs differing by a rotation in the picture plane (Fig. 2A) and for the pairs differing by a rotation in depth (Fig. 2B). In both cases, reaction time is a strikingly linear function of the angular difference between the two three-dimensional objects portrayed. The mean reaction times for individual subjects increased from a value of about 1 second at 0° of rotation for all subjects to values ranging from 4 to 6 seconds at 180° of rotation, depending upon the particular individual. Moreover, despite such variations in slope, the linearity of the function is clearly evident when the data are plotted separately for individual three-dimensional objects or for individual subjects. Polynomial regression lines were computed separately for each subject under each type of rotation. In all 16 cases the functions were found to have a highly significant linear component ($P < .001$) when tested against deviations from linearity. No significant quadratic or higher-order effects were found ($P > .05$, in all cases).

The angle through which different three-dimensional shapes must be rotated to achieve congruence is not, of course, defined. Therefore, a function like those plotted in Fig. 2 cannot be constructed in any straightforward man-

ner for the "different" pairs. The overall mean reaction time for these pairs was found, however, to be 3.8 seconds—nearly a second longer than the corresponding overall means for the "same" pairs. (In the postexperimental interview, the subjects typically reported that they attempted to rotate one end of one object into congruence with the corresponding end of the other object; they discovered that the two objects were *different* when, after this "rotation," the two free ends still remained noncongruent.)

Not only are the two functions shown in Fig. 2 both linear but they are very similar to each other with respect to intercept and slope. Indeed, for the larger angular differences the reaction times were, if anything, somewhat shorter for rotation in depth than for rotation in the picture plane. However, since this small difference is either absent or reversed in four of the eight subjects, it is of doubtful significance. The determination of identity of shape may therefore be based, in both cases, upon a process of the same general kind. If we can describe this process as some sort of "mental rotation in three-dimensional space," then the slope of the obtained functions indicates that the average rate at which these particular objects can be thus "rotated" is roughly 60° per second.

Of course the plotted reaction times necessarily include any times taken by the subjects to decide how to process

the pictures in each presented pair as well as the time taken actually to carry out the process, once it was chosen. However, even for these highly practiced subjects, the reaction times were still linear and were no more than 20 percent lower in the "pure" blocks of presentations (in which the subjects knew both the axis and the direction of the required rotation in advance of each presentation) than in the "mixed" blocks (in which the axis of rotation was unpredictable). Tentatively, this suggests that 80 percent of a typical one of these reaction times may represent some such process as "mental rotation" itself, rather than a preliminary process of preparation or search. Nevertheless, in further research now underway, we are seeking clarification of this point and others.

ROGER N. SHEPARD
JACQUELINE METZLER

Department of Psychology,
Stanford University,
Stanford, California 94305

References and Notes

1. Mrs. Jih-Jie Chang of the Bell Telephone Laboratories generated the 180 perspective projections for us by means of the Bell Laboratories' Stromberg-Carlson 4020 microfilm recorder and the computer program for constructing such projections developed there by A. M. Noll. See, for example, A. M. Noll, *Computers Automation* 14, 20 (1965).
2. We thank Mrs. Chang [see (1)]; and we also thank Dr. J. D. Elashoff for her suggestions concerning the statistical analyses. Assistance in the computer graphics was provided by the Bell Telephone Laboratories. Supported by NSF grant GS-2283 to R.N.S.

9 March 1970; revised 8 September 1970 ■

Neural Pathways Associated with Hypothalamically Elicited Attack Behavior in Cats

Abstract. Small electrolytic lesions were made in cats through electrodes, which, when stimulated, elicited either quiet biting attack or affective paw strike attack upon rats. The Nauta method for impregnating degenerating axoplasm was used to reveal that degeneration resulting from lesions at quiet attack sites followed largely along the course of the medial forebrain bundle, while the degeneration after lesions of affective attack sites was concentrated more heavily in the periventricular system.

Although it is now firmly established that the hypothalamus is intimately involved in the elaboration of aggressive behavior (1), very little is known about the neural pathways through which such behavior is mediated. In an attempt to trace out the circuits which may be associated with a cat's attack upon a rat we have employed neuroanatomic techniques in conjunction with stimulation experiments.

The development of silver stains

capable of selectively impregnating degenerating axoplasm by Nauta (2) has dramatically increased the ability of neuroanatomists to determine the polarity of conduction and the areas of termination of finely myelinated and unmyelinated fiber systems. The first step in tracing out degeneration by this technique consists in destroying a small amount of neural tissue in a selected anatomic target area and then permitting the animal to survive for a

Adjustment of Homosexual and Heterosexual Women

Adjustment of Homosexual and Heterosexual Women

By MARVIN SIEGELMAN

The traditional psychiatric belief that homosexual women are emotionally unstable (3, 10, 21) has been challenged by Armon (1), Freedman (11), and Hopkins (13). The contention that such women are neurotic has typically been voiced by clinicians reporting on their own therapy patients (3, 10, 19). One exception is the recent psychometric investigation by Kenyon (15) who studied a non-clinical group of English homosexual women, and concluded that they were higher in neuroticism than a comparison group of heterosexuals. In contrast to the 'illness' notion of homosexuality, the authors of three psychometric studies (1, 11, 13) dealing with non-clinical homosexuals and heterosexuals reported that heterosexual women were not better adjusted than homosexuals. The paucity of research in this area is exemplified by the fact that a total of only four studies, noted above, have been found to date by this author. Even the clinical literature, which is replete with case studies and therapeutic discussions concerning male homosexuality, is strikingly sparse in the area of lesbianism (19). The present study was therefore conducted to add to the small body of data we now have on the adjustment of homosexual versus heterosexual women.

METHOD

Subjects

The 84 homosexual women in the present study were not a random or representative sample, but they did not represent a clearly clinical group, such as psychiatric patients. Forty-six homosexuals were members of the New York branch of the Daughters of Bilitis (D.O.B.) organization, formed by and for lesbians. The officers of D.O.B. sent out the author's questionnaire to the D.O.B. members, and they returned them anonymously to the author; of 75 questionnaires sent out, 46 were

returned. The remaining 38 homosexuals returned questionnaires they requested from the manager of a homophile bookstore in Greenwich Village, who assisted the author; a letter had been placed in the window of the bookstore asking for volunteers. Fifty subjects requested the forms, and 38 were returned. The percentage of return was 61.3 for the D.O.B. members, 76 for the Village bookstore, and 67.2 for both sources. The occupational status for the majority of the homosexuals could be classified as professional; in addition to 11 teachers from all educational levels, and 13 graduate and undergraduate college students, there were registered nurses, social workers, editors, a statistician, a librarian, a psychiatrist, a psychologist, etc. The combining of the D.O.B. and the bookstore subjects was justified in that *t* test comparisons between the two groups for all adjustment variables, to be reported on below, showed no significant differences. The 84 homosexuals described themselves as either exclusively homosexual (70.2 per cent) or predominantly homosexual with some heterosexual tendencies (29.8 per cent). Eighty-seven per cent of the homosexuals were not in therapy. Their mean age was 30.3 (S.D. = 9.15), and they had completed an average of 15.5 years of education (S.D. = 2.47). The mean education of their parents was 11.8 years (S.D. = 4.25) for fathers and 12.1 years (S.D. = 3.21) for mothers.

The 133 heterosexual subjects consisted of undergraduate ($N = 49$) and graduate ($N = 84$) students at The City College of New York, who described themselves as exclusively heterosexual. The graduates and undergraduates represented all areas of concentration. The authors administered the questionnaires to this group during a regular class period. Ninety-three per cent of the heterosexuals were not in therapy. The mean age was 25.0 (S.D. = 7.44),

and they had completed an average of 16·5 years of education (S.D. = 4·81). The average education of their parents was 12·0 years (S.D. = 4·01) for fathers, and 11·9 years (S.D. = 6·36) for mothers.

The homosexuals did not differ significantly from the heterosexuals on education ($t = 1·66$), education of father ($t = .48$), education of mother ($t = .26$), and Schachter (17) sibling position code ($t = .48$). The two groups were different in age ($t = 4·68$, $p < .01$).

Instruments

In order to avoid a too narrow conception of adjustment, several instruments were used, which measured 12 dimensions related to mental health. The Scheier and Cattell (18) Neuroticism Scale Questionnaire (NSQ) examined four factors, Tendermindedness (I), Depression (F), Submission (E), and Anxiety (Anx.), as well as overall Neuroticism (Total NSQ). Eight additional scales measured Alienation and Trust (20), Goal-Directedness, Self-Acceptance, and Sense of Self (7), Dependency (5), Nurturance (12), and Neuroticism (16). The Crowne and Marlow (6) Social Desirability Scale (SDS) was also included in order to measure the possible contamination of giving socially desirable responses on the various adjustment measures. The correlations between the SDS and the adjustment variables for the homosexuals (noted first) and the heterosexuals (noted second) were: Tendermindedness = -·06, -·08; Depressed = ·23, ·15; Submissive = ·27, ·21; Anxiety = -·27, -·35; Total NSQ = ·10, -·03; Alienation = -·01, -·14; Trust = ·19, ·20; Goal-Directedness = ·31, ·12; Self-Accepting = ·17, -·01; Sense of Self = ·35, -·07; Dependence = -·33, -·31; Nurturance = ·17, -·07; Neuroticism = -·28, -·25. On the SDS the homosexuals indicated less of a tendency to give socially desirable responses than the heterosexuals (homosexual M = 12·3, S.D. = 5·53; heterosexual M = 14·2, S.D. = 5·65; $t = 2·47$, $p < .05$).

Analysis of covariance was used to control for age and SDS responses. In addition, partial correlations, with age and SDS responses partialled out, were computed to indicate the

amount of association between the variables and the groups (4).

RESULTS

Table I indicates that the homosexuals, in contrast to the heterosexuals, scored lower on Depression and Total NSQ, and they were higher on Goal-Direction and Self-Acceptance. The analysis of covariance results were essentially the same as the t test findings. Dignan (7) included Goal-Directed and Self-Acceptance in her 'ego-Identity' dimension based on the construct devised by Erikson (8). The association between groups and adjustment variables, as reflected in the partial correlations, were not impressive, accounting at best for only 9 per cent of the variance.

In addition to controlling for age by partial correlation and covariance statistics, the possible influence of age was estimated by calculating product-moment correlations between age and each adjustment factor. For the homosexuals there was a positive correlation between age and Alienation ($r = .22$) and Depression ($r = .22$), and a negative correlation with Dependency ($r = -.19$). There was a positive relationship, for heterosexuals, between age and Depression ($r = .28$) and Total NSQ ($r = .18$), and a negative association between age and Neuroticism ($r = -.21$). Age thus accounted for at most about 9 per cent of the variance in any adjustment factor, which would probably not effect the results very much.

NSQ data for a group of 393 presumably heterosexual females, comparable in age ($M = 31·0$) to the homosexual females reported in the present study ($M = 30·3$), was presented by Scheier and Cattell (18). The NSQ data noted for these 393 females was Tendermindedness $M = 12·2$, S.D. = 2·5; Depressed $M = 9·8$, S.D. = 2·8; Submissive $M = 13·3$, S.D. = 3·1; Anxiety $M = 9·8$, S.D. = 3·4; Total NSQ $M = 45·1$, S.D. = 7·1. The 84 homosexuals in the current research had higher scores on Tenderminded ($t = 2·21$, $p < .01$) and Dominance ($t = 5·45$, $p < .001$), and lower scores on Depression ($t = 2·15$, $p < .01$), than the 393 heterosexuals in the Scheier and Cattell report. These groups were not different on Anxiety and Total NSQ. In comparison to the

TABLE I
Comparisons of adjustment variable scores between homosexual and heterosexual females

Variables	N ^a		M		S.D.		t	P ^b	$r_{12 \cdot 3^c}$
	Homo- sexual	Hetero- sexual	Homo- sexual	Hetero- sexual	Homo- sexual	Hetero- sexual			
NSQ									
Tenderminded	82	133	13·07	13·71	2·78	2·59	1·72	.2·10	.11
Depressed	82	133	8·84	8·92	3·31	3·40	1·16	.06	.04
Submissive	82	133	10·62	12·62	3·57	3·25	4·21***	15·78***	.24***
Anxious	82	133	10·80	11·43	3·74	3·53	1·23	1·33	.08
Total NSQ	82	133	43·46	46·67	7·39	7·50	3·06*	7·48**	.19**
Ner-NSQ									
Alienated	84	130	14·48	14·23	4·87	4·66	.37	.40	.07
Trusting	84	130	15·46	15·65	4·90	4·75	.28	.14	.00
Goal-directed	84	130	9·71	32·35	6·28	6·02	5·10***	25·79***	.30***
Self-accepting	84	130	23·29	20·80	4·19	4·63	3·98***	15·27***	.23***
Sense of self	84	130	17·81	17·25	3·84	3·37	1·12	.83	.07
Dependent	84	130	15·70	16·49	5·61	5·95	.97	.12	.05
Nurturant	84	130	18·19	17·88	3·74	3·49	.62	2·26	.09
Neurotic	84	133	15·39	16·61	4·98	5·33	1·68	2·92	.11

* N differed for some variables because a few Ss did not complete all scales.

^a Analysis of covariance.

^b Partial correlation.

* .05, two-tailed.

** .01, two-tailed.

*** .001, two-tailed.

NSQ scores reported by Scheier and Cattell (18) for 315 neurotics, the 84 homosexuals in the current study had higher scores on Tendermindedness, Depression, Submission, and Anxiety, and they were not different on the Total NSQ factor.

DISCUSSION

When interpreting the results of the present study, one must of course recognize that both samples were select and clearly not representative of all heterosexual, or all homosexual, women. The data could be generalized, at best, only to women similar to those examined in the current research.

The failure to find female homosexuals more neurotic than female heterosexuals in the present study agrees with the reports of Armon (1), Freedman (11), and Hopkins (13). The NSQ findings in the current investigation were strikingly similar to the Cattell 16 PF data reported by Hopkins (13). The NSQ is part of the 16 PF and so the two can be directly compared. Hopkins noted, in agreement with the

present study, that her lesbian group was not different from her heterosexual sample on Tendermindedness, Depression, and two of the three factors comprising the Anxiety dimension, and that the lesbians were more dominant. On the Anxiety Factor differentiating the groups, the heterosexuals were more tense and excitable than the lesbians. Hopkins did not report a Total NSQ score. The willingness of the homosexuals to acknowledge deviation by joining a homophile organization (i.e. 23 of the 24 lesbians in the Hopkins study were members of the Minorities Research Group, and 46 of the 84 homosexuals in the present study belonged to the Daughters of Bilitis), and to volunteer information about themselves, may account in part for their high dominance and low anxiety scores.

The indications from the present data—weak in magnitude as they are—that the lesbians are better adjusted in some respects than the heterosexuals, has also been reported by Freedman (11) and by Hopkins (13). The psychoanalytic contention that female homosexuals have high

dependency needs (10) was not supported by Armon (1) nor by the present study, and strong independent tendencies have been found by Hopkins (13). Socardes (19) and Caprio (3) stated that female homosexuals had a deep sense of inferiority, but the results of the present study, as well as the findings of Freedman (11) and Hopkins (13) directly contradict this contention.

Both Freedman (11) in the United States and Kenyon (14) in England used the Eysenck Maudsley Personality Inventory (9), but their findings were conflicting. Although Kenyon found lesbians to score higher on Neuroticism ($M = 23.15$, S.D. = 12.28) than the heterosexual women ($M = 17.41$, S.D. = 9.6), his mean homosexual Neuroticism score was lower than the mean score of a mixed group of English university students ($M = 26.78$, S.D. = 9.28), and much lower than the average score of neurotics ($M = 31.95$, S.D. = 9.15), reported by Eysenck (9). Similarly, Kenyon did not indicate what average score was obtained by neurotics on the second test he used, the Cornell Medical Index Health Questionnaire (2) in contrast to his lesbian sample. Furthermore, the homosexual subjects in the Kenyon study included 43 women (34.9 per cent) who were either predominantly heterosexual (8.9 per cent), or bisexual (6.5 per cent) or more homosexual than heterosexual, but maintained a fair amount of heterosexual activity (19.5 per cent). These 43 women were rated between 1 and 4 on the Kinsey Scale by Kenyon (14). It was thus misleading to categorize the entire sample of 123 women as homosexual.

The need to sample a wider segment of the homosexual community is emphasized by the contradictory findings that emerge when contrasting samples are evaluated. An absence of differences in homosexual versus heterosexual adjustment is reported in most studies using non-clinical subjects. On the other hand, the writing of psychiatrists describing their patients, typically reflects the belief that homosexuals are seriously maladjusted. Although the findings of the present investigation refute the opinion that female homosexuals are more neurotic than female heterosexuals, additional studies with larger, more representative, samples, and diverse

methodologies must be conducted for a more adequate evaluation of this result.

SUMMARY

The adjustment of non-clinical homosexual women ($N = 84$) compared to heterosexuals ($N = 133$) was evaluated with the Scheier and Cattell Neuroticism Scale Questionnaire (NSQ) and eight additional factors selected from five scales. The NSQ results for homosexuals were also contrasted to the NSQ data from heterosexual women and neurotic women reported by Scheier and Cattell. The homosexuals were found to be as well adjusted as the heterosexuals. These findings were compared to four other psychometric studies involving non-clinical subjects.

ACKNOWLEDGEMENTS

The author appreciates the assistance of the Daughters of Bilitis officers and the manager of the Oscar Wilde Bookstore in securing homosexual subjects. The data processing was made possible by the free computer time supplied by the Computation Centers of The City College of New York and The University of London.

REFERENCES

1. ARMON, V. (1960). 'Some personality variables in overt female homosexuality.' *Journal of Protective Techniques*, 24, 292-309.
2. BRODMAN, K., ERDMANN, A. J., LORGE, I., GERHARDSON, C. P., and WOLFF, H. G. (1952). 'The Cornell Medical Index Health Questionnaire III: The evaluation of emotional disturbance.' *Journal of Clinical Psychology*, 8, 119-24.
3. CAPRIO, F. S. (1954). *Female Homosexuality*. New York: The Citadel Press.
4. COHEN, J. (1965). 'Some statistical issues in psychological research.' In: *Handbook of Clinical Psychology* (ed. Wolman). New York: McGraw-Hill.
5. COMRY, A. L. (1964). 'Personality factors compulsion, dependence, hostility, and neuroticism.' *Educational and Psychological Measurement*, 24, 75-84.
6. CROWNE, D. P., and MARLOWE, D. (1960). 'A new scale of social desirability independent of psychopathology.' *Journal of Consulting Psychology*, 24, 349-54.
7. DIGNAN, M. H. (1965). 'Ego identity and maternal identification.' *Journal of Personality and Social Psychology*, 1, 476-83.
8. ERIKSON, E. H. (1955). 'Growth and crises of the "healthy personality".' In: *Personality in Nature, Society, and Culture* (ed. Kluckhohn and Murray). New York: Knopf.
9. EYSENCK, H. J. (1962). *Manual of the Maudsley Personality Inventory*. San Diego, Cal.: Educational and Industrial Testing Service.
10. FENICHEL, D. (1945). *The Psychoanalytic Theory of Neurosis*. New York: Norton.
11. FREEDMAN, M. J. (1968). 'Homosexuality among women and psychological adjustment.' *Dissertation Abstracts*, 28, 2924B.
12. HARVEY, O. J., PRATHER, M. S., WHITE, J., and ALTER, R. D. (1966). 'Teachers' belief systems and preschool atmospheres.' *Journal of Educational Psychology*, 57, 373-81.
13. HOPKINS, JUNI, H. (1969). 'The lesbian personality.' *British Journal of Psychiatry*, 115, 1433-6.
14. KENYON, F. E. (1968). 'Studies in female homosexuality—Psychological test results.' *Journal of Consulting and Clinical Psychology*, 32, 510-13.
15. KENYON, F. E. (1969). 'Studies in female homosexuality: IV. Social and psychiatric aspects.' *British Journal of Psychiatry*, 114, 1337-50.
16. MC GUIRE, RUTH G. (1966). 'An inquiry into attitudes and value systems of a minority group.' Unpublished doctoral dissertation, New York University.
17. SCHACHTER, S. (1959). *The Psychology of Affiliation*. Stanford, Cal.: Stanford University Press.
18. SCHEIER, I. H., and CATTELL, R. B. (1961). *The Neuroticism Scale Questionnaire*. Champaign, Ill.: Institute for Personality and Ability Testing.
19. SOCARIDES, C. W. (1968). *The Overt Homosexual*. New York: Grune and Stratton.
20. STRUENING, E. I., and RICHARDSON, A. H. (1965). 'A factor analytic exploration of the alienation, anomie, and authoritarian domain.' *American Sociological Review*, 30, 768-78.
21. WILDER, G. B. (1965). 'Clinical aspects of female homosexuality.' In: *Sexual Inversion* (ed. Marmor). New York: Basic Books.

BY MARVIN SIEGELMAN

481

Synopsis of this paper was published in the January 1972 *Journal*.

Marvin Siegelman, Ph.D., Professor, City College of New York, New York City, 10031

Received 23 July 1971

On being sane in insane places

On Being Sane In Insane Places

David L. Rosenhan*

How do we know precisely what constitutes “normality” or mental illness? Conventional wisdom suggests that specially trained professionals have the ability to make reasonably accurate diagnoses. In this research, however, David Rosenhan provides evidence to challenge this assumption. What is -- or is not -- “normal” may have much to do with the labels that are applied to people in particular settings.

If sanity and insanity exist, how shall we know them?

The question is neither capricious nor itself insane. However much we may be personally convinced that we can tell the normal from the abnormal, the evidence is simply not compelling. It is commonplace, for example, to read about murder trials wherein eminent psychiatrists for the defense are contradicted by equally eminent psychiatrists for the prosecution on the matter of the defendant's sanity. More generally, there are a great deal of conflicting data on the reliability, utility, and meaning of such terms as “sanity,” “insanity,” “mental illness,” and “schizophrenia.” Finally, as early as 1934, {Ruth} Benedict suggested that normality and abnormality are not universal.[\[1\]](#) What is viewed as normal in one culture may be seen as quite aberrant in another. Thus, notions of normality and abnormality may not be quite as accurate as people believe they are.

To raise questions regarding normality and abnormality is in no way to question the fact that some behaviors are deviant or odd. Murder is deviant. So, too, are hallucinations. Nor does raising such questions deny the existence of the personal anguish that is often associated with “mental illness.” Anxiety and depression exist. Psychological suffering exists. But normality and abnormality, sanity and insanity, and the diagnoses that flow from them may be less substantive than many believe them to be.

At its heart, the question of whether the sane can be distinguished from the insane (and whether degrees of insanity can be distinguished from each other) is a simple matter: Do the salient characteristics that lead to diagnoses reside in the patients themselves or in the environments and contexts in which observers find them? From Bleuler, through Kretschmer, through the formulators of the recently revised Diagnostic and Statistical Manual of the American Psychiatric Association, the belief has been strong that patients present symptoms, that those symptoms can be categorized, and, implicitly, that the sane are distinguishable from the insane. More recently, however, this belief has been questioned. Based in part on theoretical and anthropological considerations, but also on philosophical, legal, and therapeutic ones, the view has grown that psychological categorization of mental illness is useless at best and downright harmful, misleading, and pejorative at worst. Psychiatric diagnoses, in this view, are in the minds of observers and are not valid summaries of characteristics displayed by the observed.

Gains can be made in deciding which of these is more nearly accurate by getting normal people (that is, people who do not have, and have never suffered, symptoms of serious psychiatric disorders) admitted to psychiatric hospitals and then determining whether they were discovered to be sane and, if so, how. If the sanity of such pseudopatients were always detected, there would be prima

facie evidence that a sane individual can be distinguished from the insane context in which he is found. Normality (and presumably abnormality) is distinct enough that it can be recognized wherever it occurs, for it is carried within the person. If, on the other hand, the sanity of the pseudopatients were never discovered, serious difficulties would arise for those who support traditional modes of psychiatric diagnosis. Given that the hospital staff was not incompetent, that the pseudopatient had been behaving as sanely as he had been out of the hospital, and that it had never been previously suggested that he belonged in a psychiatric hospital, such an unlikely outcome would support the view that psychiatric diagnosis betrays little about the patient but much about the environment in which an observer finds him.

This article describes such an experiment. Eight sane people gained secret admission to 12 different hospitals. Their diagnostic experiences constitute the data of the first part of this article; the remainder is devoted to a description of their experiences in psychiatric institutions. Too few psychiatrists and psychologists, even those who have worked in such hospitals, know what the experience is like. They rarely talk about it with former patients, perhaps because they distrust information coming from the previously insane. Those who have worked in psychiatric hospitals are likely to have adapted so thoroughly to the settings that they are insensitive to the impact of that experience. And while there have been occasional reports of researchers who submitted themselves to psychiatric hospitalization, these researchers have commonly remained in the hospitals for short periods of time, often with the knowledge of the hospital staff. It is difficult to know the extent to which they were treated like patients or like research colleagues. Nevertheless, their reports about the inside of the psychiatric hospital have been valuable. This article extends those efforts.

PSEUDOPATIENTS AND THEIR SETTINGS

The eight pseudopatients were a varied group. One was a psychology graduate student in his 20's. The remaining seven were older and "established." Among them were three psychologists, a pediatrician, a psychiatrist, a painter, and a housewife. Three pseudopatients were women, five were men. All of them employed pseudonyms, lest their alleged diagnoses embarrass them later. Those who were in mental health professions alleged another occupation in order to avoid the special attentions that might be accorded by staff, as a matter of courtesy or caution, to ailing colleagues.[\[2\]](#) With the exception myself (I was the first pseudopatient and my presence was known to the hospital administration and chief psychologist and, so far as I can tell, to them alone), the presence of pseudopatients and the nature of the research program was not known to the hospital staffs.[\[3\]](#)

The settings are similarly varied. In order to generalize the findings, admission into a variety of hospitals was sought. The 12 hospitals in the sample were located in five different states on the East and West coasts. Some were old and shabby, some were quite new. Some had good staff-patient ratios, others were quite understaffed. Only one was a strict private hospital. All of the others were supported by state or federal funds or, in one instance, by university funds.

After calling the hospital for an appointment, the pseudopatient arrived at the admissions office complaining that he had been hearing voices. Asked what the voices said, he replied that they were often unclear, but as far as he could tell they said "empty," "hollow," and "thud." The voices were unfamiliar and were of the same sex as the pseudopatient. The choice of these symptoms was occasioned by their apparent similarity to existential symptoms. Such symptoms are alleged to arise

from painful concerns about the perceived meaninglessness of one's life. It is as if the hallucinating person were saying, "My life is empty and hollow." The choice of these symptoms was also determined by the absence of a single report of existential psychoses in the literature.

Beyond alleging the symptoms and falsifying name, vocation, and employment, no further alterations of person, history, or circumstances were made. The significant events of the pseudopatient's life history were presented as they had actually occurred. Relationships with parents and siblings, with spouse and children, with people at work and in school, consistent with the aforementioned exceptions, were described as they were or had been. Frustrations and upsets were described along with joys and satisfactions. These facts are important to remember. If anything, they strongly biased the subsequent results in favor of detecting insanity, since none of their histories or current behaviors were seriously pathological in any way.

Immediately upon admission to the psychiatric ward, the pseudopatient ceased simulating any symptoms of abnormality. In some cases, there was a brief period of mild nervousness and anxiety, since none of the pseudopatients really believed that they would be admitted so easily. Indeed, their shared fear was that they would be immediately exposed as frauds and greatly embarrassed. Moreover, many of them had never visited a psychiatric ward; even those who had, nevertheless had some genuine fears about what might happen to them. Their nervousness, then, was quite appropriate to the novelty of the hospital setting, and it abated rapidly.

Apart from that short-lived nervousness, the pseudopatient behaved on the ward as he "normally" behaved. The pseudopatient spoke to patients and staff as he might ordinarily. Because there is uncommonly little to do on a psychiatric ward, he attempted to engage others in conversation. When asked by staff how he was feeling, he indicated that he was fine, that he no longer experienced symptoms. He responded to instructions from attendants, to calls for medication (which was not swallowed), and to dining-hall instructions. Beyond such activities as were available to him on the admissions ward, he spent his time writing down his observations about the ward, its patients, and the staff. Initially these notes were written "secretly," but as it soon became clear that no one much cared, they were subsequently written on standard tablets of paper in such public places as the dayroom. No secret was made of these activities.

The pseudopatient, very much as a true psychiatric patient, entered a hospital with no foreknowledge of when he would be discharged. Each was told that he would have to get out by his own devices, essentially by convincing the staff that he was sane. The psychological stresses associated with hospitalization were considerable, and all but one of the pseudopatients desired to be discharged almost immediately after being admitted. They were, therefore, motivated not only to behave sanely, but to be paragons of cooperation. That their behavior was in no way disruptive is confirmed by nursing reports, which have been obtained on most of the patients. These reports uniformly indicate that the patients were "friendly," "cooperative," and "exhibited no abnormal indications."

THE NORMAL ARE NOT DETECTABLY SANE

Despite their public "show" of sanity, the pseudopatients were never detected. Admitted, except in one case, with a diagnosis of schizophrenia,[\[4\]](#) each was discharged with a diagnosis of

schizophrenia “in remission.” The label “in remission” should in no way be dismissed as a formality, for at no time during any hospitalization had any question been raised about any pseudopatient’s simulation. Nor are there any indications in the hospital records that the pseudopatient’s status was suspect. Rather, the evidence is strong that, once labeled schizophrenic, the pseudopatient was stuck with that label. If the pseudopatient was to be discharged, he must naturally be “in remission”; but he was not sane, nor, in the institution’s view, had he ever been sane.

The uniform failure to recognize sanity cannot be attributed to the quality of the hospitals, for, although there were considerable variations among them, several are considered excellent. Nor can it be alleged that there was simply not enough time to observe the pseudopatients. Length of hospitalization ranged from 7 to 52 days, with an average of 19 days. The pseudopatients were not, in fact, carefully observed, but this failure speaks more to traditions within psychiatric hospitals than to lack of opportunity.

Finally, it cannot be said that the failure to recognize the pseudopatients’ sanity was due to the fact that they were not behaving sanely. While there was clearly some tension present in all of them, their daily visitors could detect no serious behavioral consequences—nor, indeed, could other patients. It was quite common for the patients to “detect” the pseudopatient’s sanity. During the first three hospitalizations, when accurate counts were kept, 35 of a total of 118 patients on the admissions ward voiced their suspicions, some vigorously. “You’re not crazy. You’re a journalist, or a professor (referring to the continual note-taking). You’re checking up on the hospital.” While most of the patients were reassured by the pseudopatient’s insistence that he had been sick before he came in but was fine now, some continued to believe that the pseudopatient was sane throughout his hospitalization. The fact that the patients often recognized normality when staff did not raises important questions.

Failure to detect sanity during the course of hospitalization may be due to the fact that physicians operate with a strong bias toward what statisticians call the Type 2 error. This is to say that physicians are more inclined to call a healthy person sick (a false positive, Type 2) than a sick person healthy (a false negative, Type 1). The reasons for this are not hard to find: it is clearly more dangerous to misdiagnose illness than health. Better to err on the side of caution, to suspect illness even among the healthy.

But what holds for medicine does not hold equally well for psychiatry. Medical illnesses, while unfortunate, are not commonly pejorative. Psychiatric diagnoses, on the contrary, carry with them personal, legal, and social stigmas. It was therefore important to see whether the tendency toward diagnosing the sane insane could be reversed. The following experiment was arranged at a research and teaching hospital whose staff had heard these findings but doubted that such an error could occur in their hospital. The staff was informed that at some time during the following three months, one or more pseudopatients would attempt to be admitted into the psychiatric hospital. Each staff member was asked to rate each patient who presented himself at admissions or on the ward according to the likelihood that the patient was a pseudopatient. A 10-point scale was used, with a 1 and 2 reflecting high confidence that the patient was a pseudopatient.

Judgments were obtained on 193 patients who were admitted for psychiatric treatment. All staff who had had sustained contact with or primary responsibility for the patient – attendants, nurses,

psychiatrists, physicians, and psychologists – were asked to make judgments. Forty-one patients were alleged, with high confidence, to be pseudopatients by at least one member of the staff. Twenty-three were considered suspect by at least one psychiatrist. Nineteen were suspected by one psychiatrist and one other staff member. Actually, no genuine pseudopatient (at least from my group) presented himself during this period.

The experiment is instructive. It indicates that the tendency to designate sane people as insane can be reversed when the stakes (in this case, prestige and diagnostic acumen) are high. But what can be said of the 19 people who were suspected of being “sane” by one psychiatrist and another staff member? Were these people truly “sane” or was it rather the case that in the course of avoiding the Type 2 error the staff tended to make more errors of the first sort – calling the crazy “sane”? There is no way of knowing. But one thing is certain: any diagnostic process that lends itself too readily to massive errors of this sort cannot be a very reliable one.

THE STICKINESS OF PSYCHODIAGNOSTIC LABELS

Beyond the tendency to call the healthy sick – a tendency that accounts better for diagnostic behavior on admission than it does for such behavior after a lengthy period of exposure – the data speak to the massive role of labeling in psychiatric assessment. Having once been labeled schizophrenic, there is nothing the pseudopatient can do to overcome the tag. The tag profoundly colors others’ perceptions of him and his behavior.

From one viewpoint, these data are hardly surprising, for it has long been known that elements are given meaning by the context in which they occur. Gestalt psychology made the point vigorously, and Asch^[5] demonstrated that there are “central” personality traits (such as “warm” versus “cold”) which are so powerful that they markedly color the meaning of other information in forming an impression of a given personality. “Insane,” “schizophrenic,” “manic-depressive,” and “crazy” are probably among the most powerful of such central traits. Once a person is designated abnormal, all of his other behaviors and characteristics are colored by that label. Indeed, that label is so powerful that many of the pseudopatients’ normal behaviors were overlooked entirely or profoundly misinterpreted. Some examples may clarify this issue.

Earlier, I indicated that there were no changes in the pseudopatient’s personal history and current status beyond those of name, employment, and, where necessary, vocation. Otherwise, a veridical description of personal history and circumstances was offered. Those circumstances were not psychotic. How were they made consonant with the diagnosis modified in such a way as to bring them into accord with the circumstances of the pseudopatient’s life, as described by him?

As far as I can determine, diagnoses were in no way affected by the relative health of the circumstances of a pseudopatient’s life. Rather, the reverse occurred: the perception of his circumstances was shaped entirely by the diagnosis. A clear example of such translation is found in the case of a pseudopatient who had had a close relationship with his mother but was rather remote from his father during his early childhood. During adolescence and beyond, however, his father became a close friend, while his relationship with his mother cooled. His present relationship with his wife was characteristically close and warm. Apart from occasional angry exchanges, friction was minimal. The children had rarely been spanked. Surely there is nothing especially pathological about

such a history. Indeed, many readers may see a similar pattern in their own experiences, with no markedly deleterious consequences. Observe, however, how such a history was translated in the psychopathological context, this from the case summary prepared after the patient was discharged.

This white 39-year-old male . . . manifests a long history of considerable ambivalence in close relationships, which begins in early childhood. A warm relationship with his mother cools during his adolescence. A distant relationship with his father is described as becoming very intense. Affective stability is absent. His attempts to control emotionality with his wife and children are punctuated by angry outbursts and, in the case of the children, spankings. And while he says that he has several good friends, one senses considerable ambivalence embedded in those relationships also . . .

The facts of the case were unintentionally distorted by the staff to achieve consistency with a popular theory of the dynamics of a schizophrenic reaction. Nothing of an ambivalent nature had been described in relations with parents, spouse, or friends. To the extent that ambivalence could be inferred, it was probably not greater than is found in all human's relationships. It is true the pseudopatient's relationships with his parents changed over time, but in the ordinary context that would hardly be remarkable – indeed, it might very well be expected. Clearly, the meaning ascribed to his verbalizations (that is, ambivalence, affective instability) was determined by the diagnosis: schizophrenia. An entirely different meaning would have been ascribed if it were known that the man was “normal.”

All pseudopatients took extensive notes publicly. Under ordinary circumstances, such behavior would have raised questions in the minds of observers, as, in fact, it did among patients. Indeed, it seemed so certain that the notes would elicit suspicion that elaborate precautions were taken to remove them from the ward each day. But the precautions proved needless. The closest any staff member came to questioning those notes occurred when one pseudopatient asked his physician what kind of medication he was receiving and began to write down the response. “You needn’t write it,” he was told gently. “If you have trouble remembering, just ask me again.”

If no questions were asked of the pseudopatients, how was their writing interpreted? Nursing records for three patients indicate that the writing was seen as an aspect of their pathological behavior. “Patient engaged in writing behavior” was the daily nursing comment on one of the pseudopatients who was never questioned about his writing. Given that the patient is in the hospital, he must be psychologically disturbed. And given that he is disturbed, continuous writing must be behavioral manifestation of that disturbance, perhaps a subset of the compulsive behaviors that are sometimes correlated with schizophrenia.

One tacit characteristic of psychiatric diagnosis is that it locates the sources of aberration within the individual and only rarely within the complex of stimuli that surrounds him. Consequently, behaviors that are stimulated by the environment are commonly misattributed to the patient's disorder. For example, one kindly nurse found a pseudopatient pacing the long hospital corridors. “Nervous, Mr. X?” she asked. “No, bored,” he said.

The notes kept by pseudopatients are full of patient behaviors that were misinterpreted by well-intentioned staff. Often enough, a patient would go “berserk” because he had, wittingly or

unwittingly, been mistreated by, say, an attendant. A nurse coming upon the scene would rarely inquire even cursorily into the environmental stimuli of the patient's behavior. Rather, she assumed that his upset derived from his pathology, not from his present interactions with other staff members. Occasionally, the staff might assume that the patient's family (especially when they had recently visited) or other patients had stimulated the outburst. But never were the staff found to assume that one of themselves or the structure of the hospital had anything to do with a patient's behavior. One psychiatrist pointed to a group of patients who were sitting outside the cafeteria entrance half an hour before lunchtime. To a group of young residents he indicated that such behavior was characteristic of the oral-acquisitive nature of the syndrome. It seemed not to occur to him that there were very few things to anticipate in a psychiatric hospital besides eating.

A psychiatric label has a life and an influence of its own. Once the impression has been formed that the patient is schizophrenic, the expectation is that he will continue to be schizophrenic. When a sufficient amount of time has passed, during which the patient has done nothing bizarre, he is considered to be in remission and available for discharge. But the label endures beyond discharge, with the unconfirmed expectation that he will behave as a schizophrenic again. Such labels, conferred by mental health professionals, are as influential on the patient as they are on his relatives and friends, and it should not surprise anyone that the diagnosis acts on all of them as a self-fulfilling prophecy. Eventually, the patient himself accepts the diagnosis, with all of its surplus meanings and expectations, and behaves accordingly.

The inferences to be made from these matters are quite simple. Much as Zigler and Phillips have demonstrated that there is enormous overlap in the symptoms presented by patients who have been variously diagnosed,[\[6\]](#) so there is enormous overlap in the behaviors of the sane and the insane. The sane are not "sane" all of the time. We lose our tempers "for no good reason." We are occasionally depressed or anxious, again for no good reason. And we may find it difficult to get along with one or another person – again for no reason that we can specify. Similarly, the insane are not always insane. Indeed, it was the impression of the pseudopatients while living with them that they were sane for long periods of time – that the bizarre behaviors upon which their diagnoses were allegedly predicated constituted only a small fraction of their total behavior. If it makes no sense to label ourselves permanently depressed on the basis of an occasional depression, then it takes better evidence than is presently available to label all patients insane or schizophrenic on the basis of bizarre behaviors or cognitions. It seems more useful, as Mischel[\[7\]](#) has pointed out, to limit our discussions to behaviors the stimuli that provoke them, and their correlates.

It is not known why powerful impressions of personality traits, such as "crazy" or "insane," arise. Conceivably, when the origins of and stimuli that give rise to a behavior are remote or unknown, or when the behavior strikes us as immutable, trait labels regarding the behavior arise. When, on the other hand, the origins and stimuli are known and available, discourse is limited to the behavior itself. Thus, I may hallucinate because I am sleeping, or I may hallucinate because I have ingested a peculiar drug. These are termed sleep-induced hallucinations, or dreams, and drug-induced hallucinations, respectively. But when the stimuli to my hallucinations are unknown, that is called craziness, or schizophrenia –as if that inference were somehow as illuminating as the others.

THE EXPERIENCE OF PSYCHIATRIC HOSPITALIZATION

The term “mental illness” is of recent origin. It was coined by people who were humane in their inclinations and who wanted very much to raise the station of (and the public’s sympathies toward) the psychologically disturbed from that of witches and “crazies” to one that was akin to the physically ill. And they were at least partially successful, for the treatment of the mentally ill has improved considerably over the years. But while treatment has improved, it is doubtful that people really regard the mentally ill in the same way that they view the physically ill. A broken leg is something one recovers from, but mental illness allegedly endures forever. A broken leg does not threaten the observer, but a crazy schizophrenic? There is by now a host of evidence that attitudes toward the mentally ill are characterized by fear, hostility, aloofness, suspicion, and dread. The mentally ill are society’s lepers.

That such attitudes infect the general population is perhaps not surprising, only upsetting. But that they affect the professionals – attendants, nurses, physicians, psychologists and social workers – who treat and deal with the mentally ill is more disconcerting, both because such attitudes are self-evidently pernicious and because they are unwitting. Most mental health professionals would insist that they are sympathetic toward the mentally ill, that they are neither avoidant nor hostile. But it is more likely that an exquisite ambivalence characterizes their relations with psychiatric patients, such that their avowed impulses are only part of their entire attitude. Negative attitudes are there too and can easily be detected. Such attitudes should not surprise us. They are the natural offspring of the labels patients wear and the places in which they are found.

Consider the structure of the typical psychiatric hospital. Staff and patients are strictly segregated. Staff have their own living space, including their dining facilities, bathrooms, and assembly places. The glassed quarters that contain the professional staff, which the pseudopatients came to call “the cage,” sit out on every dayroom. The staff emerge primarily for care-taking purposes – to give medication, to conduct therapy or group meeting, to instruct or reprimand a patient. Otherwise, staff keep to themselves, almost as if the disorder that afflicts their charges is somehow catching.

So much is patient-staff segregation the rule that, for four public hospitals in which an attempt was made to measure the degree to which staff and patients mingle, it was necessary to use “time out of the staff cage” as the operational measure. While it was not the case that all time spent out of the cage was spent mingling with patients (attendants, for example, would occasionally emerge to watch television in the dayroom), it was the only way in which one could gather reliable data on time for measuring.

The average amount of time spent by attendants outside of the cage was 11.3 percent (range, 3 to 52 percent). This figure does not represent only time spent mingling with patients, but also includes time spent on such chores as folding laundry, supervising patients while they shave, directing ward cleanup, and sending patients to off-ward activities. It was the relatively rare attendant who spent time talking with patients or playing games with them. It proved impossible to obtain a “percent mingling time” for nurses, since the amount of time they spent out of the cage was too brief. Rather, we counted instances of emergence from the cage. On the average, daytime nurses emerged from the cage 11.5 times per shift, including instances when they left the ward entirely (range, 4 to 39 times). Later afternoon and night nurses were even less available, emerging on the average 9.4 times per shift (range, 4 to 41 times). Data on early morning nurses, who arrived usually

after midnight and departed at 8 a.m., are not available because patients were asleep during most of this period.

Physicians, especially psychiatrists, were even less available. They were rarely seen on the wards. Quite commonly, they would be seen only when they arrived and departed, with the remaining time being spent in their offices or in the cage. On the average, physicians emerged on the ward 6.7 times per day (range, 1 to 17 times). It proved difficult to make an accurate estimate in this regard, since physicians often maintained hours that allowed them to come and go at different times.

The hierarchical organization of the psychiatric hospital has been commented on before, but the latent meaning of that kind of organization is worth noting again. Those with the most power have the least to do with patients, and those with the least power are the most involved with them. Recall, however, that the acquisition of role-appropriate behaviors occurs mainly through the observation of others, with the most powerful having the most influence. Consequently, it is understandable that attendants not only spend more time with patients than do any other members of the staff – that is required by their station in the hierarchy – but, also, insofar as they learn from their superior's behavior, spend as little time with patients as they can. Attendants are seen mainly in the cage, which is where the models, the action, and the power are.

I turn now to a different set of studies, these dealing with staff response to patient-initiated contact. It has long been known that the amount of time a person spends with you can be an index of your significance to him. If he initiates and maintains eye contact, there is reason to believe that he is considering your requests and needs. If he pauses to chat or actually stops and talks, there is added reason to infer that he is individuating you. In four hospitals, the pseudopatients approached the staff member with a request which took the following form: "Pardon me, Mr. [or Dr. or Mrs.] X, could you tell me when I will be eligible for grounds privileges?" (or "... when I will be presented at the staff meeting?" or "... when I am likely to be discharged?"). While the content of the question varied according to the appropriateness of the target and the pseudopatient's (apparent) current needs the form was always a courteous and relevant request for information. Care was taken never to approach a particular member of the staff more than once a day, lest the staff member become suspicious or irritated . . . [R]emember that the behavior of the pseudopatients was neither bizarre nor disruptive. One could indeed engage in good conversation with them.

. . . Minor differences between these four institutions were overwhelmed by the degree to which staff avoided continuing contacts that patients had initiated. By far, their most common response consisted of either a brief response to the question, offered while they were "on the move" and with head averted, or no response at all. The encounter frequently took the following bizarre form: (pseudopatient) "Pardon me, Dr. X. Could you tell me when I am eligible for grounds privileges?" (physician) "Good morning, Dave. How are you today? (Moves off without waiting for a response.) . . .

POWERLESSNESS AND DEPERSONALIZATION

Eye contact and verbal contact reflect concern and individuation; their absence, avoidance and depersonalization. The data I have presented do not do justice to the rich daily encounters that grew up around matters of depersonalization and avoidance. I have records of patients who were beaten by

staff for the sin of having initiated verbal contact. During my own experience, for example, one patient was beaten in the presence of other patients for having approached an attendant and told him, "I like you." Occasionally, punishment meted out to patients for misdemeanors seemed so excessive that it could not be justified by the most rational interpretations of psychiatric cannon. Nevertheless, they appeared to go unquestioned. Tempers were often short. A patient who had not heard a call for medication would be roundly excoriated, and the morning attendants would often wake patients with, "Come on, you m_____ f_____, out of bed!"

Neither anecdotal nor "hard" data can convey the overwhelming sense of powerlessness which invades the individual as he is continually exposed to the depersonalization of the psychiatric hospital. It hardly matters which psychiatric hospital – the excellent public ones and the very plush private hospital were better than the rural and shabby ones in this regard, but, again, the features that psychiatric hospitals had in common overwhelmed by far their apparent differences.

Powerlessness was evident everywhere.

The patient is deprived of many of his legal rights by dint of his psychiatric commitment. He is shorn of credibility by virtue of his psychiatric label. His freedom of movement is restricted. He cannot initiate contact with the staff, but may only respond to such overtures as they make. Personal privacy is minimal. Patient quarters and possessions can be entered and examined by any staff member, for whatever reason. His personal history and anguish is available to any staff member (often including the "grey lady" and "candy striper" volunteer) who chooses to read his folder, regardless of their therapeutic relationship to him. His personal hygiene and waste evacuation are often monitored. The water closets have no doors.

At times, depersonalization reached such proportions that pseudopatients had the sense that they were invisible, or at least unworthy of account. Upon being admitted, I and other pseudopatients took the initial physical examinations in a semipublic room, where staff members went about their own business as if we were not there.

On the ward, attendants delivered verbal and occasionally serious physical abuse to patients in the presence of others (the pseudopatients) who were writing it all down. Abusive behavior, on the other hand, terminated quite abruptly when other staff members were known to be coming. Staff are credible witnesses. Patients are not.

A nurse unbuttoned her uniform to adjust her brassiere in the present of an entire ward of viewing men. One did not have the sense that she was being seductive. Rather, she didn't notice us. A group of staff persons might point to a patient in the dayroom and discuss him animatedly, as if he were not there.

One illuminating instance of depersonalization and invisibility occurred with regard to medication. All told, the pseudopatients were administered nearly 2100 pills, including Elavil, Stelazine, Compazine, and Thorazine, to name but a few. (That such a variety of medications should have been administered to patients presenting identical symptoms is itself worthy of note.) Only two were swallowed. The rest were either pocketed or deposited in the toilet. The pseudopatients were not alone in this. Although I have no precise records on how many patients rejected their medications, the pseudopatients frequently found the medications of other patients in the toilet before

they deposited their own. As long as they were cooperative, their behavior and the pseudopatients' own in this matter, as in other important matters, went unnoticed throughout.

Reactions to such depersonalization among pseudopatients were intense. Although they had come to the hospital as participant observers and were fully aware that they did not "belong," they nevertheless found themselves caught up in and fighting the process of depersonalization. Some examples: a graduate student in psychology asked his wife to bring his textbooks to the hospital so he could "catch up on his homework" – this despite the elaborate precautions taken to conceal his professional association. The same student, who had trained for quite some time to get into the hospital, and who had looked forward to the experience, "remembered" some drag races that he had wanted to see on the weekend and insisted that he be discharged by that time. Another pseudopatient attempted a romance with a nurse. Subsequently, he informed the staff that he was applying for admission to graduate school in psychology and was very likely to be admitted, since a graduate professor was one of his regular hospital visitors. The same person began to engage in psychotherapy with other patients – all of this as a way of becoming a person in an impersonal environment.

THE SOURCES OF DEPERSONALIZATION

What are the origins of depersonalization? I have already mentioned two. First are attitudes held by all of us toward the mentally ill – including those who treat them – attitudes characterized by fear, distrust, and horrible expectations on the one hand, and benevolent intentions on the other. Our ambivalence leads, in this instance as in others, to avoidance.

Second, and not entirely separate, the hierarchical structure of the psychiatric hospital facilitates depersonalization. Those who are at the top have least to do with patients, and their behavior inspires the rest of the staff. Average daily contact with psychiatrists, psychologists, residents, and physicians combined ranged from 3.9 to 25.1 minutes, with an overall mean of 6.8 (six pseudopatients over a total of 129 days of hospitalization). Included in this average are time spent in the admissions interview, ward meetings in the presence of a senior staff member, group and individual psychotherapy contacts, case presentation conferences and discharge meetings. Clearly, patients do not spend much time in interpersonal contact with doctoral staff. And doctoral staff serve as models for nurses and attendants.

There are probably other sources. Psychiatric installations are presently in serious financial straits. Staff shortages are pervasive, and that shortens patient contact. Yet, while financial stresses are realities, too much can be made of them. I have the impression that the psychological forces that result in depersonalization are much stronger than the fiscal ones and that the addition of more staff would not correspondingly improve patient care in this regard. The incidence of staff meetings and the enormous amount of record-keeping on patients, for example, have not been as substantially reduced as has patient contact. Priorities exist, even during hard times. Patient contact is not a significant priority in the traditional psychiatric hospital, and fiscal pressures do not account for this. Avoidance and depersonalization may.

Heavy reliance upon psychotropic medication tacitly contributes to depersonalization by convincing staff that treatment is indeed being conducted and that further patient contact may not be necessary. Even here, however, caution needs to be exercised in understanding the role of

psychotropic drugs. If patients were powerful rather than powerless, if they were viewed as interesting individuals rather than diagnostic entities, if they were socially significant rather than social lepers, if their anguish truly and wholly compelled our sympathies and concerns, would we not seek contact with them, despite the availability of medications? Perhaps for the pleasure of it all?

THE CONSEQUENCES OF LABELING AND DEPERSONALIZATION

Whenever the ratio of what is known to what needs to be known approaches zero, we tend to invent “knowledge” and assume that we understand more than we actually do. We seem unable to acknowledge that we simply don’t know. The needs for diagnosis and remediation of behavioral and emotional problems are enormous. But rather than acknowledge that we are just embarking on understanding, we continue to label patients “schizophrenic,” “manic-depressive,” and “insane,” as if in those words we captured the essence of understanding. The facts of the matter are that we have known for a long time that diagnoses are often not useful or reliable, but we have nevertheless continued to use them. We now know that we cannot distinguish sanity from insanity. It is depressing to consider how that information will be used.

Not merely depressing, but frightening. How many people, one wonders, are sane but not recognized as such in our psychiatric institutions? How many have been needlessly stripped of their privileges of citizenship, from the right to vote and drive to that of handling their own accounts? How many have feigned insanity in order to avoid the criminal consequences of their behavior, and, conversely, how many would rather stand trial than live interminably in a psychiatric hospital – but are wrongly thought to be mentally ill? How many have been stigmatized by well-intentioned, but nevertheless erroneous, diagnoses? On the last point, recall again that a “Type 2 error” in psychiatric diagnosis does not have the same consequences it does in medical diagnosis. A diagnosis of cancer that has been found to be in error is cause for celebration. But psychiatric diagnoses are rarely found to be in error. The label sticks, a mark of inadequacy forever.

Finally, how many patients might be “sane” outside the psychiatric hospital but seem insane in it – not because craziness resides in them, as it were, but because they are responding to a bizarre setting, one that may be unique to institutions which harbor neither people? Goffman [8] calls the process of socialization to such institutions “mortification” – an apt metaphor that includes the processes of depersonalization that have been described here. And while it is impossible to know whether the pseudopatients’ responses to these processes are characteristic of all inmates – they were, after all, not real patients – it is difficult to believe that these processes of socialization to a psychiatric hospital provide useful attitudes or habits of response for living in the “real world.”

SUMMARY AND CONCLUSIONS

It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals. The hospital itself imposes a special environment in which the meaning of behavior can easily be misunderstood. The consequences to patients hospitalized in such an environment – the powerlessness, depersonalization, segregation, mortification, and self-labeling – seem undoubtedly counter-therapeutic.

I do not, even now, understand this problem well enough to perceive solutions. But two matters seem to have some promise. The first concerns the proliferation of community mental health

facilities, of crisis intervention centers, of the human potential movement, and of behavior therapies that, for all of their own problems, tend to avoid psychiatric labels, to focus on specific problems and behaviors, and to retain the individual in a relatively non-pejorative environment. Clearly, to the extent that we refrain from sending the distressed to insane places, our impressions of them are less likely to be distorted. (The risk of distorted perceptions, it seems to me, is always present, since we are much more sensitive to an individual's behaviors and verbalizations than we are to the subtle contextual stimuli than often promote them. At issue here is a matter of magnitude. And, as I have shown, the magnitude of distortion is exceedingly high in the extreme context that is a psychiatric hospital.)

The second matter that might prove promising speaks to the need to increase the sensitivity of mental health workers and researchers to the Catch 22 position of psychiatric patients. Simply reading materials in this area will be of help to some such workers and researchers. For others, directly experiencing the impact of psychiatric hospitalization will be of enormous use. Clearly, further research into the social psychology of such total institutions will both facilitate treatment and deepen understanding.

I and the other pseudopatients in the psychiatric setting had distinctly negative reactions. We do not pretend to describe the subjective experiences of true patients. Theirs may be different from ours, particularly with the passage of time and the necessary process of adaptation to one's environment. But we can and do speak to the relatively more objective indices of treatment within the hospital. It could be a mistake, and a very unfortunate one, to consider that what happened to us derived from malice or stupidity on the part of the staff. Quite the contrary, our overwhelming impression of them was of people who really cared, who were committed and who were uncommonly intelligent. Where they failed, as they sometimes did painfully, it would be more accurate to attribute those failures to the environment in which they, too, found themselves than to personal callousness. Their perceptions and behaviors were controlled by the situation, rather than being motivated by a malicious disposition. In a more benign environment, one that was less attached to global diagnosis, their behaviors and judgments might have been more benign and effective.

* I thank W. Mischel, E. Orne, and M.S. Rosenhan for comments on an earlier draft of this manuscript.

SOURCE: David L. Rosenhan, "On Being Sane in Insane Places," *Science*, Vol. 179 (Jan. 1973), 250-258.

Copyright 1973 by the American Association for the Advancement of Science.

[1] R. Benedict, J.Gen. Psychol., 10 (1934), 59.

[2] Beyond the personal difficulties that the pseudopatient is likely to experience in the hospital, there are legal and social ones that, combined, require considerable attention before entry. For example, once admitted to a psychiatric institution, it is difficult, if not impossible, to be discharged on short notice, state law to the contrary notwithstanding. I was not sensitive to these difficulties at the outset of the project, nor to the personal and situational emergencies that can arise, but later a writ of habeas corpus was prepared for each of the entering pseudopatients and an attorney was kept "on call" during every hospitalization. I am grateful to John Kaplan and Robert Bartels for legal advice and assistance in these matters.

[3] However distasteful such concealment is, it was a necessary first step to examining these questions. Without concealment, there would have been no way to know how valid these experiences

were; nor was there any way of knowing whether whatever detections occurred were a tribute to the diagnostic acumen of the hospital's rumor network. Obviously, since my concerns are general ones that cut across individual hospitals and staffs, I have respected their anonymity and have eliminated clues that might lead to their identification.

[4] Interestingly, of the 12 admissions, 11 were diagnosed as schizophrenic and one, with the identical symptomatology, as manic-depressive psychosis. This diagnosis has more favorable prognosis, and it was given by the private hospital in our sample. One the relations between social class and psychiatric diagnosis, see A. deB. Hollingshead and F.C. Redlich, *Social Class and Mental Illness: A Community Study* (New York: John Wiley, 1958).

[5] S.E. Asch, *J. Abnorm. Soc. Psychol.*, 41 (1946), *Social Psychology* (Englewood Cliffs, NJ: Prentice_Hall, 1952).

[6] E. Zigler and L. Phillips, *J. Abnorm. Soc. Psychol.* 63, (1961) 69. See also R. K. Freudenberg and J. P. Robertson, *A.M.A. Arch. Neurol. Psychiatr.*, 76, (1956), 14.

[7] W. Mischel, *Personality and Assessment* (New York; John Wiley, 1968).

[8] E. Goffman, *Asylums* (Garden City, NY; Doubleday, 1961).

[BACK to Leslie's Home Page](#)

Interpersonal Dynamics in a Simulated Prison

Interpersonal Dynamics in a Simulated Prison

CRAIG HANEY, CURTIS BANKS AND PHILIP ZIMBARDO
*Department of Psychology, Stanford University,
California 94305, U.S.A.*

Interpersonal dynamics in a prison environment were studied experimentally by designing a functional simulation of a prison in which subjects role-played prisoners and guards for an extended period of time. To assess the power of the social forces on the emergent behaviour in this situation, alternative explanations in terms of pre-existing dispositions were eliminated through subject selection. A homogeneous, "normal" sample was chosen after extensive interviewing and diagnostic testing of a large group of volunteer male college students. Half of the subjects were randomly assigned to role-play prison guards for eight hours each day, while the others role-played prisoners incarcerated for nearly one full week. Neither group received any specific training in these roles.

Continuous, direct observation of behavioural interactions was supplemented by video-taped recording, questionnaires, self-report scales and interviews. All these data sources converge on the conclusion that this simulated prison developed into a psychologically compelling prison environment. As such, it elicited unexpectedly intense, realistic and often pathological reactions from many of the participants. The prisoners experienced a loss of personal identity and the arbitrary control of their behaviour which resulted in a syndrome of passivity, dependency, depression and helplessness. In contrast, the guards (with rare exceptions) experienced a marked gain in social power, status and group identification which made role-playing rewarding.

The most dramatic of the coping behaviour utilised by half of the prisoners in adapting to this stressful situation was the development of acute emotional disturbance—severe enough to warrant their early release. At least a third of the guards were judged to have become far more aggressive and dehumanising toward the prisoners than would ordinarily be predicted in a simulation study. Only a very few of the observed reactions to this experience of imprisonment could be attributed to personality trait differences which existed before the subjects began to play their assigned roles.

Introduction

After he had spent four years in a Siberian prison the great Russian novelist Dostoevsky commented, surprisingly, that his time in prison had created in him a deep optimism about the ultimate future of mankind because, as he put it, if man could survive the horrors of prison life he must surely be a "creature who could withstand anything". The cruel irony which Dostoevsky overlooked is that the reality of prison bears witness not only to the resilience and adaptiveness of the men who tolerate life within its walls, but as well to the "ingenuity" and tenacity of those who devised and still maintain our correctional and reformatory systems.

Nevertheless, in the century which has passed since Dostoevsky's imprisonment, little has changed to render the main thrust of his statement less relevant. Although we have passed through periods of enlightened humanitarian reform, in which physical conditions within prisons have improved somewhat and the rhetoric of rehabilitation has replaced the language of punitive incarceration, the social institution of prison has continued to fail. On purely pragmatic grounds, there is substantial evidence that prisons in fact neither "rehabilitate" nor act as a deterrent to future crime—in America, recidivism rates upwards of 75% speak quite decisively to these criteria. And, to perpetuate what is additionally an economic failure, American taxpayers alone must provide an expenditure for "corrections" of 1.5 billion dollars annually. On humanitarian grounds as well, prisons have failed: our mass media are increasingly filled with accounts of atrocities committed daily, man against man, in reaction to the penal system or in the name of it. The experience of prison undeniably creates, almost to the point of cliché, an intense hatred and disrespect in most inmates for the authority and the established order of society into which they will eventually return. And the toll which it takes on the deterioration of human spirit for those who must administer it, as well as for those upon whom it is inflicted, is incalculable.

Attempts to provide an explanation of the deplorable condition of our penal system and its dehumanising effects upon prisoners and guards, often focus upon what might be called the *dispositional hypothesis*. While this explanation is rarely expressed explicitly, it is central to a prevalent non-conscious ideology: that the state of the social institution of prison is due to the "nature" of the people who administer it, or the "nature" of the people who populate it, or both. That is, a major contributing cause to despicable conditions, violence, brutality, dehumanisation and degradation existing within any prison can be traced to some innate or acquired characteristic of the correctional and inmate population. Thus on the one hand, there is the contention that violence and brutality exist within prison because guards are sadistic, uneducated, and insensitive people. It is the "guard mentality", a unique syndrome of negative traits which they bring into the situation, that engenders the inhumane treatment of prisoners. Or, from other quarters comes the argument that violence and brutality in prison are the logical and predictable result of the

involuntary confinement of a collective of individuals whose life histories are, by definition, characterised by disregard for law, order and social convention and a concurrent propensity for impulsiveness and aggression. Logically, it follows that these individuals, having proved themselves incapable of functioning satisfactorily within the "normal" structure of society, cannot do so either inside the structure provided by prisons. To control such men as these, the argument continues, whose basic orientation to any conflict situation is to react with physical power or deception, force must be met with force, and a certain number of violent encounters must be expected and tolerated by the public.

The dispositional hypothesis has been embraced by the proponents of the prison *status quo* (blaming conditions on the evil in the prisoners), as well as by its critics (attributing the evil to guards and staff with their evil motives and deficient personality structures). The appealing simplicity of this proposition localises the source of prison riots, recidivism and corruption in these "bad seeds" and not in the conditions of the "prison soil". Such an analysis directs attention away from the complex matrix of social, economic and political forces which combine to make prisons what they are—and which would require complex, expensive, revolutionary solutions to bring about any meaningful change. Instead, rioting prisoners are identified, punished, transferred to maximum security institutions or shot, outside agitators sought and corrupt officials suspended—while the system itself goes on essentially unchanged, its basic structure unexamined and unchallenged.

However, a critical evaluation of the dispositional hypothesis cannot be made directly through observation in existing prison settings, since such naturalistic observation necessarily confounds the acute effects of the environment with the chronic characteristics of the inmate and guard populations. To separate the effects of the prison environment *per se* from those attributable to *à priori* dispositions of its inhabitants requires a research strategy in which a "new" prison is constructed, comparable in its fundamental social-psychological milieu to existing prison systems, but entirely populated by individuals who are undifferentiated in all essential dimensions from the rest of society.

Such was the approach taken in the present empirical study, namely, to create a prison-like situation in which the guards and inmates were initially comparable and characterised as being "normal-average", and then to observe the patterns of behaviour which resulted, as well as the cognitive, emotional and attitudinal reactions which emerged. Thus, we began our experiment with a sample of individuals who did not deviate from the normal range of the general population on a variety of dimensions we were able to measure. Half were randomly assigned to the role of "prisoner", the others to that of "guard", neither group having any history of crime, emotional disability, physical handicap nor even intellectual or social disadvantage.

The environment created was that of a "mock" prison which physically constrained the prisoners in barred cells and psychologically conveyed the sense of imprisonment to all participants. Our intention was not to create a *literal*

simulation of an American prison, but rather a functional representation of one. For ethical, moral and pragmatic reasons we could not detain our subjects for extended or indefinite periods of time, we could not exercise the threat and promise of severe physical punishment, we could not allow homosexual or racist practices to flourish, nor could we duplicate certain other specific aspects of prison life. Nevertheless, we believed that we could create a situation with sufficient mundane realism to allow the role-playing participants to go beyond the superficial demands of their assignment into the deep structure of the characters they represented. To do so, we established functional equivalents for the activities and experiences of actual prison life which were expected to produce qualitatively similar psychological reactions in our subjects—feelings of power and powerlessness, of control and oppression, of satisfaction and frustration, of arbitrary rule and resistance to authority, of status and anonymity, of machismo and emasculation. In the conventional terminology of experimental social psychology, we first identified a number of relevant conceptual variables through analysis of existing prison situations, then designed a setting in which these variables were made operational. No specific hypotheses were advanced other than the general one that assignment to the treatment of “guard” or “prisoner” would result in significantly different reactions on behavioural measures of interaction, emotional measures of mood state and pathology, attitudes toward self, as well as other indices of coping and adaptation to this novel situation. What follows is the mechanics of how we created and peopled our prison, what we observed, what our subjects reported, and finally, what we can conclude about the nature of the prison environment and the experience of imprisonment which can account for the failure of our prisons.

Method

Overview

The effects of playing the role of “guard” or “prisoner” were studied in the context of an experimental simulation of a prison environment. The research design was a relatively simple one, involving as it did only a single treatment variable, the random assignment to either a “guard” or “prisoner” condition. These roles were enacted over an extended period of time (nearly one week) within an environment which was physically constructed to resemble a prison. Central to the methodology of creating and maintaining a psychological state of imprisonment was the functional simulation of significant properties of “real prison life” (established through information from former inmates, correctional personnel and texts).

The “guards” were free with certain limits to implement the procedures of induction into the prison setting and maintenance of custodial retention of the “prisoners”. These inmates, having voluntarily submitted to the conditions of this total institution in which they now lived, coped in various ways with its

stresses and its challenges. The behaviour of both groups of subjects was observed, recorded and analysed. The dependent measures were of two general types: transactions between and within each group of subjects, recorded on video and audio tape as well as directly observed; individual reactions on questionnaires, mood inventories, personality tests, daily guard shift reports, and post experimental interviews.

Subjects

The 21 subjects who participated in the experiment were selected from an initial pool of 75 respondents, who answered a newspaper advertisement asking for male volunteers to participate in a psychological study of "prison life" in return for payment of \$15 per day. Those who responded to the notice completed an extensive questionnaire concerning their family background, physical and mental health history, prior experience and attitudinal propensities with respect to sources of psychopathology (including their involvement in crime). Each respondent who completed the background questionnaire was interviewed by one of two experimenters. Finally, the 24 subjects who were judged to be most stable (physically and mentally), most mature, and least involved in anti-social behaviour were selected to participate in the study. On a random basis, half of the subjects were assigned the role of "guard", half to the role of "prisoner".

The subjects were normal, healthy males attending colleges throughout the United States who were in the Stanford area during the summer. They were largely of middle class socio-economic status, Caucasians (with the exception of one Oriental subject). Initially they were strangers to each other, a selection precaution taken to avoid the disruption of any pre-existing friendship patterns and to mitigate against any transfer into the experimental situation of previously established relationships or patterns of behaviour.

This final sample of subjects was administered a battery of psychological tests on the day prior to the start of the simulation, but to avoid any selective bias on the part of the experimenter-observers, scores were not tabulated until the study was completed.

Two subjects who were assigned to be a "stand-by" in case an additional "prisoner" was needed were not called, and one subject assigned to be a "stand-by" guard decided against participating just before the simulation phase began—thus, our data analysis is based upon ten prisoners and eleven guards in our experimental conditions.

Procedure

Physical aspects of the prison

The prison was built in a 35-ft section of a basement corridor in the psychology building at Stanford University. It was partitioned by two fabricated walls, one of which was fitted with the only entrance door to the cell block, the other

contained a small observation screen. Three small cells (6 x 9 ft) were made from converted laboratory rooms by replacing the usual doors with steel barred, black painted ones, and removing all furniture.

A cot (with mattress, sheet and pillow) for each prisoner was the only furniture in the cells. A small closet across from the cells served as a solitary confinement facility; its dimensions were extremely small (2 x 2 x 7 ft) and it was unlit.

In addition, several rooms in an adjacent wing of the building were used as guards' quarters (to change in and out of uniform or for rest and relaxation), a bedroom for the "warden" and "superintendent", and an interview-testing room. Behind the observation screen at one end of the "yard" was video recording equipment and sufficient space for several observers.

Operational details

The "prisoner" subjects remained in the mock-prison 24 hours per day for the duration of the study. Three were arbitrarily assigned to each of the three cells; the others were on stand-by call at their homes. The "guard" subjects worked on three-man, eight-hour shifts; remaining in the prison environment only during their work shift, going about their usual lives at other times.

Role instruction

All subjects had been told that they would be assigned either the guard or the prisoner role on a completely random basis and all had voluntarily agreed to play either role for \$15.00 per day for up to two weeks. They signed a contract guaranteeing a minimally adequate diet, clothing, housing and medical care as well as the financial remuneration in return for their stated "intention" of serving in the assigned role for the duration of the study.

It was made explicit in the contract that those assigned to be prisoners should expect to be under surveillance (have little or no privacy) and to have some of their basic civil rights suspended during their imprisonment, excluding physical abuse. They were given no other information about what to expect nor instructions about behaviour appropriate for a prisoner role. Those actually assigned to this treatment were informed by phone to be available at their place of residence on a given Sunday when we would start the experiment.

The subjects assigned to be guards attended an orientation meeting on the day prior to the induction of the prisoners. At this time they were introduced to the principal investigators, the "Superintendent" of the prison (P.G.Z.) and an undergraduate research assistant who assumed the administrative role of "Warden". They were told that we wanted to try to simulate a prison environment within the limits imposed by pragmatic and ethical considerations. Their assigned task was to "maintain the reasonable degree of order within the prison necessary for its effective functioning", although the specifics of how this

duty might be implemented were not explicitly detailed. They were made aware of the fact that while many of the contingencies with which they might be confronted were essentially unpredictable (e.g. prisoner escape attempts), part of their task was to be prepared for such eventualities and to be able to deal appropriately with the variety of situations that might arise. The "Warden" instructed the guards in the administrative details, including: the work-shifts, the mandatory daily completion of shift reports concerning the activity of guards and prisoners, the completion of "critical incident" reports which detailed unusual occurrences and the administration of meals, work and recreation programmes for the prisoners. In order to begin to involve these subjects in their roles even before the first prisoner was incarcerated, the guards assisted in the final phases of completing the prison complex—putting the cots in the cells, signs on the walls, setting up the guards' quarters, moving furniture, water coolers, refrigerators, etc.

The guards generally believed that we were primarily interested in studying the behaviour of the prisoners. Of course, we were equally interested in the effect which enacting the role of guard in this environment would have on their behaviour and subjective states.

To optimise the extent to which their behaviour would reflect their genuine reactions to the experimental prison situation and not simply their ability to follow instructions, they were intentionally given only minimal guidelines for what it meant to be a guard. An explicit and categorical prohibition against the use of physical punishment or physical aggression was, however, emphasised by the experimenters. Thus, with this single notable exception, their roles were relatively unstructured initially, requiring each "guard" to carry out activities necessary for interacting with a group of "prisoners" as well as with other "guards" and the "correctional staff".

Uniform

In order to promote feelings of anonymity in the subjects each group was issued identical uniforms. For the guards, the uniform consisted of: plain khaki shirts and trousers, a whistle, a police night stick (wooden batons) and reflecting sunglasses which made eye contact impossible. The prisoners' uniform consisted of loosely fitting muslin smocks with an identification number on front and back. No underclothes were worn beneath these "dresses". A chain and lock were placed around one ankle. On their feet they wore rubber sandals and their hair was covered with a nylon stocking made into a cap. Each prisoner was also issued a toothbrush, soap, soapdish, towel and bed linen. No personal belongings were allowed in the cells.

The outfitting of both prisoners and guards in this manner served to enhance group identity and reduce individual uniqueness within the two groups. The khaki uniforms were intended to convey a military attitude, while the whistle and night-stick were carried as symbols of control and power. The prisoners'

uniforms were designed not only to deindividuate the prisoners but to be humiliating and serve as symbols of their dependence and subservience. The ankle chain was a constant reminder (even during their sleep when it hit the other ankle) of the oppressiveness of the environment. The stocking cap removed any distinctiveness associated with hair length, colour or style (as does shaving of heads in some "real" prisons and the military). The ill-fitting uniforms made the prisoners feel awkward in their movements; since these dresses were worn without undergarments, the uniforms forced them to assume unfamiliar postures, more like those of a woman than a man—another part of the emasculating process of becoming a prisoner.

Induction procedure

With the cooperation of Palo Alto City Police Department all of the subjects assigned to the prisoner treatment were unexpectedly "arrested" at their residences. A police officer charged them with suspicion of burglary or armed robbery, advised them of their legal rights, handcuffed them, thoroughly searched them (often as curious neighbours looked on) and carried them off to the police station in the rear of the police car. At the station they went through the standard routines of being fingerprinted, having an identification file prepared and then being placed in a detention cell. Each prisoner was blindfolded and subsequently driven by one of the experimenters and a subject-guard to our mock prison. Throughout the entire arrest procedure, the police officers involved maintained a formal, serious attitude, avoiding answering any questions of clarification as to the relation of this "arrest" to the mock prison study.

Upon arrival at our experimental prison, each prisoner was stripped, sprayed with a delousing preparation (a deodorant spray) and made to stand alone naked for a while in the cell yard. After being given the uniform described previously and having an I.D. picture taken ("mug shot"), the prisoner was put in his cell and ordered to remain silent.

Administrative routine

When all the cells were occupied, the warden greeted the prisoners and read them the rules of the institution (developed by the guards and the warden). They were to be memorised and to be followed. Prisoners were to be referred to only by the number on their uniforms, also in an effort to depersonalise them.

The prisoners were to be served three bland meals per day, were allowed three supervised toilet visits, and given two hours daily for the privilege of reading or letterwriting. Work assignments were issued for which the prisoners were to receive an hourly wage to constitute their \$15 daily payment. Two visiting periods per week were scheduled, as were movie rights and exercise periods. Three times a day all prisoners were lined up for a "count" (one on each guard

work-shift). The initial purpose of the "count" was to ascertain that all prisoners were present, and to test them on their knowledge of the rules and their I.D. numbers. The first perfunctory counts lasted only about 10 minutes, but on each successive day (or night) they were spontaneously increased in duration until some lasted several hours. Many of the pre-established features of administrative routine were modified or abandoned by the guards, and some were forgotten by the staff over the course of the study.

Data collection (dependent measures)

The exploratory nature of this investigation and the absence of specific hypotheses led us to adopt the strategy of surveying as many as possible behavioural and psychological manifestations of the prison experience on the guards and the prisoners. In fact, one major methodological problem in a study of this kind is defining the limits of the "data", since relevant data emerged from virtually every interaction between any of the participants, as well as from subjective and behavioural reactions of individual prisoners, guards, the warden, superintendent, research assistants and visitors to the prison. It will also be clear when the results are presented that causal direction cannot always be established in the patterns of interaction where any given behaviour might be the consequence of a current or prior instigation by another subject and, in turn, might serve as impetus for eliciting reactions from others.

Data collection was organised around the following sources:

(1) *Videotaping.* About 12 hours of recordings were made of daily, regularly occurring events, such as the counts and meals, as well as unusual interactions, such as a prisoner rebellion, visits from a priest, a lawyer and parents, Parole Board meetings and others. Concealed video equipment recorded these events through a screen in the partition at one end of the cell-block yard or in a conference room (for parole meetings).

(2) *Audio recording.* Over 30 hours of recordings were made of verbal interactions between guards and prisoners on the prison yard. Concealed microphones picked up all conversation taking place in the yard as well as some within the cells. Other concealed recordings were made in the testing-interview room on selected occasions—interactions between the warden, superintendent and the prisoners' Grievance Committee, parents, other visitors and prisoners released early. In addition, each subject was interviewed by one of the experimenters (or by other research associates) during the study, and most just prior to its termination.

(3) *Rating scales.* Mood adjective checklists and sociometric measures were administered on several occasions to assess emotional changes in affective state and interpersonal dynamics among the guard and prisoner groups.

(4) *Individual difference scales.* One day prior to the start of the simulation all subjects completed a series of paper and pencil personality tests. These tests

were selected to provide dispositional indicators of interpersonal behaviour styles—the *F* scale of Authoritarian Personality [1], and the Machiavellianism Scale [2]—as well as areas of possible personality pathology through the newly developed Comrey Personality Scale [3]. The subscales of this latter test consist of:

- (a) trustworthiness
- (b) orderliness
- (c) conformity
- (d) activity
- (e) stability
- (f) extroversion
- (g) masculinity
- (h) empathy

(5) *Personal observations.* The guards made daily reports of their observations after each shift, the experimenters kept informal diaries and all subjects completed post-experimental questionnaires of their reactions to the experience about a month after the study was over.

Data analyses presented problems of several kinds. First, some of the data was subject to possible errors due to selective sampling. The video and audio recordings tended to be focussed upon the more interesting, dramatic events which occurred. Over time, the experimenters became more personally involved in the transaction and were not as distant and objective as they should have been. Second, there are not complete data on all subjects for each measure because of prisoners being released at different times and because of unexpected disruptions, conflicts and administrative problems. Finally, we have a relatively small sample on which to make cross-tabulations by possible independent and individual difference variables.

However, despite these shortcomings some of the overall effects in the data are powerful enough to reveal clear, reliable results. Also some of the more subtle analyses were able to yield statistically significant results even with the small sample size. Most crucial for the conclusions generated by this exploratory study is the consistency in the pattern of relationships which emerge across a wide range of measuring instruments and different observers. Special analyses were required only of the video and audio material, the other data sources were analysed following established scoring procedures.

Video analysis

There were 25 relatively discrete incidents identifiable on the tapes of prisoner-guard interactions. Each incident or scene was scored for the presence of nine behavioural (and verbal) categories. Two judges who had not been involved with the simulation study scored these tapes. These categories were defined as follows:

Question. All questions asked, requests for information or assistance (excluding rhetorical questions).

Command. An order to commence or abstain from a specific behaviour, directed either to individuals or groups. Also generalised orders, e.g. "Settle down".

Information. A specific piece of information proffered by anyone whether requested or not, dealing with any contingency of the simulation.

Individualizing reference. Positive: use of a person's real name, nickname or allusion to special positive physical characteristics. Negative: use of prison number, title, generalised "you" or reference to derogatory characteristic.

Threat. Verbal statement of contingent negative consequences of a wide variety, e.g. no meal, long count, pushups, lock-up in hole, no visitors, etc.

Deprecation insult. Use of obscenity, slander, malicious statement directed toward individual or group, e.g. "You lead a life of mendacity" or "You guys are really stupid."

Resistance. Any physical resistance, usually prisoners to guards, such as holding on to beds, blocking doors, shoving guard or prisoner, taking off stocking caps, refusing to carry out orders.

Help. Person physically assisting another (i.e. excludes verbal statements of support), e.g. guard helping another to open door, prisoner helping another prisoner in cleanup duties.

Use of instruments. Use of any physical instrument to either intimidate, threaten, or achieve specific end, e.g. fire extinguisher, batons, whistles.

Audio analysis

For purposes of classifying the verbal behaviour recorded from interviews with guards and prisoners, eleven categories were devised. Each statement made by the interviewee was assigned to the appropriate category by judges. At the end of this process for any given interview analysis, a list had been compiled of the nature and frequencies of the interviewee's discourse. The eleven categories for assignment of verbal expressions were:

Questions. All questions asked, requests for information or assistance (excluding rhetorical questions).

Informative statements. A specific piece of information proffered by anyone whether requested or not, dealing with any contingency of the simulation.

Demands. Declarative statements of need or imperative requests.

Requests. Deferential statements for material or personal consideration.

Commands. Orders to commence or abstain from a specific behaviour, directed either to individuals or groups.

Outlook, positive/negative. Expressions of expectancies for future experiences or future events; either negative or positive in tone, e.g. "I don't think I can make it" v. "I believe I will feel better."

Criticism. Expressions of critical evaluation concerning other subjects, the experimenters or the experiment itself.

Statements of identifying reference, deindividuating/individuating. Statements wherein a subject makes some reference to another subject specifically by allusion to given name or distinctive characteristics (individuating reference), or by allusion to non-specific identity or institutional number (deindividuating reference).

Desire to continue. Any expression of a subject's wish to continue or to curtail participation in the experiment.

Self-evaluation, positive/negative. Statements of self-esteem or self-degradation, e.g. "I feel pretty good about the way I've adjusted" v. "I hate myself for being so oppressive."

Action intentions, positive/negative including "intent to aggress". Statements concerning interviewees' intentions to do something in the future, either of a positive, constructive nature or a negative, destructive nature, e.g. "I'm not going to be so mean from now on" v. "I'll break the door down."

Results

Overview

Although it is difficult to anticipate exactly what the influence of incarceration will be upon the individuals who are subjected to it and those charged with its maintenance (especially in a simulated reproduction), the results of the present experiment support many commonly held conceptions of prison life and validate anecdotal evidence supplied by articulate ex-convicts. The environment of arbitrary custody had great impact upon the affective states of both guards and prisoners as well as upon the interpersonal processes taking place between and within those role-groups.

In general, guards and prisoners showed a marked tendency toward increased negativity of affect and their overall outlook became increasingly negative. As the experiment progressed, prisoners expressed intentions to do harm to others more frequently. For both prisoners and guards, self-evaluations were more deprecating as the experience of the prison environment became internalised.

Overt behaviour was generally consistent with the subjective self-reports and affective expressions of the subjects. Despite the fact that guards and prisoners were essentially free to engage in any form of interaction (positive or negative, supportive or affrontive, etc.), the characteristic nature of their encounters tended to be negative, hostile, affrontive and dehumanising. Prisoners immediately adopted a generally passive response mode while guards assumed a very active initiating role in all interactions. Throughout the experiment, commands were the most frequent form of verbal behaviour and, generally, verbal exchanges were strikingly impersonal, with few references to individual identity. Although it was clear to all subjects that the experimenters would not

permit physical violence to take place, varieties of less direct aggressive behaviour were observed frequently (especially on the part of guards). In lieu of physical violence, verbal affronts were used as one of the most frequent forms of interpersonal contact between guards and prisoners.

The most dramatic evidence of the impact of this situation upon the participants was seen in the gross reactions of five prisoners who had to be released because of extreme emotional depression, crying, rage and acute anxiety. The pattern of symptoms was quite similar in four of the subjects and began as early as the second day of imprisonment. The fifth subject was released after being treated for a psychosomatic rash which covered portions of his body. Of the remaining prisoners, only two said they were not willing to forfeit the money they had earned in return for being "paroled". When the experiment was terminated prematurely after only six days, all the remaining prisoners were delighted by their unexpected good fortune. In contrast, most of the guards seemed to be distressed by the decision to stop the experiment and it appeared to us that had become sufficiently involved in their roles so that they now enjoyed the extreme control and power which they exercised and were reluctant to give it up. One guard did report being personally upset at the suffering of the prisoners and claimed to have considered asking to change his role to become one of them—but never did so. None of the guards ever failed to come to work on time for their shift, and indeed, on several occasions guards remained on duty voluntarily and uncomplaining for extra hours—without additional pay.

The extremely pathological reactions which emerged in both groups of subjects testify to the power of the social forces operating, but still there were individual differences seen in styles of coping with this novel experience and in degrees of successful adaptation to it. Half the prisoners did endure the oppressive atmosphere, and not all the guards resorted to hostility. Some guards were tough but fair ("played by the rules"), some went far beyond their roles to engage in creative cruelty and harassment, while a few were passive and rarely instigated any coercive control over the prisoners.

These differential reactions to the experience of imprisonment were not suggested by or predictable from the self-report measures of personality and attitude or the interviews taken before the experiment began. The standardised tests employed indicated that a perfectly normal emotionally stable sample of subjects had been selected. In those few instances where differential test scores do discriminate between subjects, there is an opportunity to, partially at least, discern some of the personality variables which may be critical in the adaptation to and tolerance of prison confinement.

Initial personality and attitude measures

Overall, it is apparent that initial personality-attitude dispositions account for an extremely small part of the variation in reactions to this mock prison experience. However, in a few select instances, such dispositions do seem to be correlated with the prisoners' ability to adjust to the experimental prison environment.

Comrey scale

The Comrey Personality Inventory [3] was the primary personality scale administered to both guards and prisoners. The mean scores for prisoners and guards on the eight sub-scales of the test are shown in Table 1. No differences between prisoner and guard mean scores on any scale even approach statistical significance. Furthermore, in no case does any group mean fall outside of the 40 to 60 centile range of the normative male population reported by Comrey.

Table 1. Mean scores for prisoners and guards on eight Comrey subscales

Scale	Prisoners	Guards
Trustworthiness—high score indicates belief in the basic honesty and good intentions of others	$\bar{X} = 92.56$	$\bar{X} = 89.64$
Orderliness—extent to which person is meticulous and concerned with neatness and orderliness	$\bar{X} = 75.67$	$\bar{X} = 73.82$
Conformity—indicates belief in law enforcement, acceptance of society as it is, resentment of nonconformity in others	$\bar{X} = 65.67$	$\bar{X} = 63.18$
Activity—liking for physical activity, hard work, and exercise	$\bar{X} = 89.78$	$\bar{X} = 91.73$
Stability—high score indicates calm, optimistic, stable, confident individual	$\bar{X} = 98.33$	$\bar{X} = 101.45$
Extroversion—suggests outgoing, easy to meet person	$\bar{X} = 83.22$	$\bar{X} = 81.91$
Masculinity—"people who are not bothered by crawling creatures, the sight of blood, vulgarity, who do not cry easily and are not interested in love stories"	$\bar{X} = 88.44$	$\bar{X} = 87.00$
Empathy—high score indicates individuals who are sympathetic, helpful, generous and interested in devoting their lives to the service of others	$\bar{X} = 91.78$	$\bar{X} = 95.36$

Table 2. Mean scores for "Remaining" v. "Early released" prisoners on Comrey subscales

Scale	Remaining prisoners	Early released prisoners	Mean difference
Trustworthiness	93.4	90.8	+2.6
Orderliness	76.6	78.0	-1.4
Conformity	67.2	59.4	+7.8
Activity	91.4	86.8	+4.6
Stability	99.2	99.6	-0.4
Extroversion	98.4	76.2	+22.2
Masculinity	91.6	86.0	+5.6
Empathy	103.8	85.6	+17.2

Table 2 shows the mean scores on the Comrey sub-scales for prisoners who remained compared with prisoners who were released early due to severe emotional reactions to the environment. Although none of the comparisons achieved statistical significance, three seemed at least suggestive as possible discriminators of those who were able to tolerate this type of confinement and those who were not. Compared with those who had to be released, prisoners who remained in prison until the termination of the study: scored higher on conformity ("acceptance of society as it is"), showed substantially higher average scores on Comrey's measure of extroversion and also scored higher on a scale of empathy (helpfulness, sympathy and generosity).

F-Scale

The *F*-scale is designed to measure rigid adherence to conventional values and a submissive, uncritical attitude towards authority. There was no difference between the mean score for prisoners (4.78) and the mean score for guards (4.36) on this scale.

Again, comparing those prisoners who remained with those who were released early, we notice an interesting trend. This intra-group comparison shows remaining prisoners scoring more than twice as high on conventionality and authoritarianism ($\bar{X} = 7.78$) than those prisoners released early ($\bar{X} = 3.20$). While the difference between these means fails to reach acceptable levels of significance, it is striking to note that a rank-ordering of prisoners on the *F*-scale correlates highly with the duration of their stay in the experiment ($r_s = 0.898$, $P < 0.005$). To the extent that a prisoner was high in rigidity, in adherence to conventional values, and in the acceptance of authority, he was likely to remain longer and adjust more effectively to this authoritarian prison environment.

Machiavellianism

There were no significant mean differences found between guards ($\bar{X} = 7.73$) and prisoners ($\bar{X} = 8.77$) on this measure of effective interpersonal manipulation. In addition, the Mach Scale was of no help in predicting the likelihood that a prisoner would tolerate the prison situation and remain in the study until its termination.

This latter finding, the lack of any mean differences between prisoners who remained *v.* those who were released from the study, is somewhat surprising since one might expect the Hi Mach's skill at manipulating social interaction and mediating favourable outcomes for himself might be acutely relevant to the simulated prison environment. Indeed, the two prisoners who scored highest on the Machiavellianism scale were also among those adjudged by the experimenters to have made unusually effective adaptations to their confinement. Yet, paradoxically (and this may give the reader some feeling for the anomalies we encountered in attempting to predict in-prison behaviour from personality

measures), the other two prisoners whom we categorised as having effectively adjusted to confinement actually obtained the lowest Mach scores of any prisoners.

Video recordings

An analysis of the video recordings indicates a preponderance of genuinely negative interactions, i.e. physical aggression, threats, deprecations, etc. It is also clear that any assertive activity was largely the prerogative of the guards, while prisoners generally assumed a relatively passive demeanour. Guards more often aggressed, more often insulted, more often threatened. Prisoners, when they reacted at all, engaged primarily in resistance to these guard behaviours.

For guards, the most frequent verbal behaviour was the giving of commands and their most frequent form of physical behaviour was aggression. The most frequent form of prisoners' verbal behaviour was question-asking, their most frequent form of physical behaviour was resistance. On the other hand, the most infrequent behaviour engaged in overall throughout the experiment was "helping"—only one such incident was noted from all the video recording collected. That solitary sign of human concern for a fellow occurred between two prisoners.

Although question-asking was the most frequent form of verbal behaviour for the prisoners, guards actually asked questions more frequently overall than did prisoners (but not significantly so). This is reflective of the fact that the overall level of behaviour emitted was much higher for the guards than for the prisoners. All of those verbal acts categorised as commands were engaged in by guards. Obviously, prisoners had no opportunity to give commands at all, that behaviour becoming the exclusive "right" of guards.

Of a total 61 incidents of direct interpersonal reference observed (incidents in which one subject spoke directly to another with the use of some identifying reference, i.e. "Hey, Peter"; "you there", etc.), 58 involved the use of some deindividuating rather than some individuating form of reference. (Recall that we characterised this distinction as follows: an individuating reference involved the use of a person's actual name, nickname or allusion to special physical characteristics, whereas a deindividuating reference involved the use of a prison number, or a generalised "you"—thus being a very depersonalising form of reference.) Since all subjects were at liberty to refer to one another in either mode, it is significant that such a large proportion of the references noted involved were in the deindividuating mode ($Z = 6.9, P < 0.01$). Deindividuating references were made more often by guards in speaking to prisoners than the reverse ($Z = 3.67, P < 0.01$). (This finding, as all prisoner-guard comparisons for specific categories, may be somewhat confounded by the fact that guards apparently enjoyed a greater freedom to initiate verbal as well as other forms of behaviour. Note, however, that the existence of this greater "freedom" on the part of the guards is itself an empirical finding since it was not prescribed

a priori.) It is of additional interest to point out that in the only three cases in which verbal exchange involved some individuating reference, it was prisoners who personalised guards.

A total of 32 incidents were observed which involved a verbal threat spoken by one subject to another. Of these, 27 such incidents involved a guard threatening a prisoner. Again, the indulgence of guards in this form of behaviour was significantly greater than the indulgence of prisoners, the observed frequencies deviating significantly from an equal distribution of threats across both groups ($Z = 3.88, P < 0.01$).

Guards more often deprecated and insulted prisoners than prisoners did of guards. Of a total of 67 observed incidents, the deprecation-insult was expressed disproportionately by guards to prisoners 61 times; ($Z = 6.72, P < 0.01$).

Physical resistance was observed 34 different times. Of these, 32 incidents involved resistance by a prisoner. Thus, as we might expect, at least in this reactive behaviour domain, prisoner responses far exceeded those of the guards ($Z = 5.14, P < 0.01$).

The use of some object or instrument in the achievement of an intended purpose or in some interpersonal interaction was observed 29 times. Twenty-three such incidents involved the use of an instrument by a guard rather than a prisoner. This disproportionate frequency is significantly variant from an equal random use by both prisoners and guards ($Z = 3.16, P < 0.01$).

Over time, from day to day, guards were observed to generally escalate their harassment of the prisoners. In particular, a comparison of two of the first prisoner-guard interactions (during the counts) with two of the last counts in the experiment yielded significant differences in: the use of deindividuating references per unit time ($\bar{X}_{t_1} = 0.0$ and $\bar{X}_{t_2} = 5.40$, respectively; $t = 3.65, P < 0.10$); the incidence of deprecation-insult per unit time ($\bar{X}_{t_1} = 0.3$ and $\bar{X}_{t_2} = 5.70$, respectively; $t = 3.16, P < 0.10$). On the other hand, a temporal analysis of the prisoner video data indicated a general decrease across all categories over time: prisoners came to initiate acts far less frequently and responded (if at all) more passively to the acts of others—they simply *behaved less*.

Although the harassment by the guards escalated overall as the experiment wore on, there was some variation in the extent to which the three different guard shifts contributed to the harassment in general. With the exception of the 2.30 a.m. count, prisoners enjoyed some respite during the late night guard shift (10.00 p.m. to 6.00 a.m.). But they really were “under the gun” during the evening shift. This was obvious in our observations and in subsequent interviews with the prisoners and was also confirmed in analysis of the video taped interactions. Comparing the three different guard shifts, the evening shift was significantly different from the other two in resorting to commands; the means being 9.30 and 4.04, respectively, for standardised units of time ($t = 2.50, P < 0.05$). In addition, the guards on this “tough and cruel” shift showed more than twice as many deprecation-insults toward the prisoners (means of 5.17 and

2.29, respectively, $P < 0.20$). They also tended to use instruments more often than other shifts to keep the prisoners in line.

Audio recordings

The audio recordings made throughout the prison simulation afforded one opportunity to systematically collect self-report data from prisoners and guards regarding (among other things) their emotional reactions, their outlook, and their interpersonal evaluations and activities within the experimental setting. Recorded interviews with both prisoners and guards offered evidence that: guards tended to express nearly as much negative outlook and negative self-regard as most prisoners (one concerned guard, in fact, expressed more negative self-regard than any prisoner and more general negative affect than all but one of the prisoners); prisoner interviews were marked by negativity in expressions of affect, self-regard and action intentions (including intent to aggress and negative outlook).

Analysis of the prisoner interviews also gave *post hoc* support to our informal impressions and subjective decisions concerning the differential emotional effects of the experiment upon those prisoners who remained and those who were released early from the study. A comparison of the mean number of expressions of negative outlook, negative affect, negative self-regard and intentions to aggress made by remaining *v.* released prisoners (per interview) yielded the following results: prisoners released early expressed more negative expectations during interviews than those who remained ($t = 2.32, P < 0.10$) and also more negative affect ($t = 2.17, P < 0.10$); prisoners released early expressed more negative self-regard, and four times as many "intentions to aggress" as prisoners who remained (although those comparisons fail to reach an acceptable level of significance).

Since we could video-record only public interactions on the "yard", it was of special interest to discover what was occurring among prisoners in private. What were they talking about in the cells—their college life, their vocation, girl friends, what they would do for the remainder of the summer once the experiment was over. We were surprised to discover that fully 90% of all conversations among prisoners were related to prison topics, while only 10% to non-prison topics such as the above. They were most concerned about food, guard harassment, setting up a grievance committee, escape plans, visitors, reactions of prisoners in the other cells and in solitary. Thus, in their private conversations when they might escape the roles they were playing in public, they did not. There was no discontinuity between their presentation of self when under surveillance and when alone.

Even more remarkable was the discovery that the prisoners had begun to adopt and accept the guards' negative attitude toward them. Half of all reported private interactions between prisoners could be classified as non-supportive and non-cooperative. Moreover, when prisoners made evaluative statements of or

expressed regard for, their fellow prisoners, 85% of the time they were uncomplimentary and deprecating. This set of observed frequencies departs significantly from chance expectations based on a conservative binomial probability frequency ($P < 0.01$ for prison *v.* non-prison topics; $P < 0.05$ for negative *v.* positive or neutral regard).

Mood adjective self-reports

Twice during the progress of the experiment each subject was asked to complete a mood adjective checklist and indicate his current affective state. The data gleaned from these self-reports did not lend themselves readily to statistical analysis. However, the trends suggested by simple enumeration are important enough to be included without reference to statistical significance. In these written self-reports, prisoners expressed nearly three times as much negative as positive affect. Prisoners roughly expressed three times as much negative affect as guards. Guards expressed slightly more negative than positive affect. While prisoners expressed about twice as much emotionality as did guards, a comparison of mood self-reports over time reveals that the prisoners showed two to three times as much mood fluctuation as did the relatively stable guards. On the dimension of activity-passivity, prisoners tended to score twice as high, indicating twice as much internal "agitation" as guards (although, as stated above, prisoners were seen to be markedly less active than guards in terms of overt behaviour).

It would seem from these results that while the experience had a categorically negative emotional impact upon both guards and prisoners, the effects upon prisoners were more profound and unstable.

When the mood scales were administered for a third time, just after the subjects were told the study had been terminated (and the early released subjects returned for the debriefing encounter session), marked changes in mood were evident. All of the now "ex-convicts" selected self-descriptive adjectives which characterised their mood as less negative and much more positive. In addition, they now felt less passive than before. There were no longer any differences on the sub-scales of this test between prisoners released early and those who remained throughout. Both groups of subjects had returned to their pre-experimental baselines of emotional responding. This seems to reflect the situational specificity of the depression and stress reactions experienced while in the role of prisoner.

Representative personal statements

Much of the flavour and impact of this prison experience is unavoidably lost in the relatively formal, objective analyses outlined in this paper. The following quotations taken from interviews, conversations and questionnaires provide a more personal view of what it was like to be a prisoner or guard in the "Stanford County Prison" experiment.

Guards

"They [the prisoners] seemed to lose touch with the reality of the experiment—they took me so seriously."

"...I didn't interfere with any of the guards' actions. Usually if what they were doing bothered me, I would walk out and take another duty."

"...looking back, I am impressed by how little I felt for them . . ."

"...They [the prisoners] didn't see it as an experiment. It was real and they were fighting to keep their identity. But we were always there to show them just who was boss."

"...I was tired of seeing the prisoners in their rags and smelling the strong odours of their bodies that filled the cells. I watched them tear at each other, on orders given by us."

"...Acting authoritatively can be fun. Power can be a great pleasure."

"...During the inspection, I went to cell 2 to mess up a bed which the prisoner had made and he grabbed me, screaming that he had just made it, and he wasn't going to let me mess it up. He grabbed my throat, and although he was laughing I was pretty scared. I lashed out with my stick and hit him in the chin (although not very hard) and when I freed myself I became angry."

Prisoners

"...The way we were made to degrade ourselves really brought us down and that's why we all sat docile towards the end of the experiment."

"...I realise now (after it's over) that no matter how together I thought I was inside my head, my prison behaviour was often less under my control than I realised. No matter how open, friendly and helpful I was with other prisoners I was still operating as an isolated, self-centred person, being rational rather than compassionate."

"...I began to feel I was losing my identity, that the person I call _____, the person who volunteered to get me into this prison (because it was a prison to me, it *still* is a prison to me, I don't regard it as an experiment or a simulation ...) was distant from me, was remote until finally I wasn't *that* person, I was 416. I was really my number and 416 was really going to have to decide what to do."

"I learned that people can easily forget that others are human."

Debriefing encounter sessions

Because of the unexpectedly intense reactions (such as the above) generated by this mock-prison experience, we decided to terminate the study at the end of six days rather than continue for the second week. Three separate encounter sessions were held, first, for the prisoners, then for the guards and finally for all participants together. Subjects and staff openly discussed their reactions and strong feelings were expressed and shared. We analysed the moral conflicts posed by this experience and used the debriefing sessions to make explicit alternative courses of action that would lead to more moral behaviour in future comparable situations.

Follow-ups on each subject over the year following termination of the study revealed the negative effects of participation had been temporary, while the personal gain to the subjects endured.

Conclusions and Discussion

It should be apparent that the elaborate procedures (and staging) employed by the experimenters to insure a high degree of mundane realism in this mock prison contributed to its effective functional simulation of the psychological dynamics operating in "real" prisons. We observed empirical relationships in the simulated prison environment which were strikingly isomorphic to the internal relations of real prisons, corroborating many of the documented reports of what occurs behind prison walls.

The conferring of differential power on the status of "guard" and "prisoner" constituted, in effect, the institutional validation of those roles. But further, many of the subjects ceased distinguishing between prison role and their prior self-identities. When this occurred, within what was a surprisingly short period of time, we witnessed a sample of normal, healthy American college students fractionate into a group of prison guards who seemed to derive pleasure from insulting, threatening, humiliating and dehumanising their peers—those who by chance selection had been assigned to the "prisoner" role. The typical prisoner syndrome was one of passivity, dependency, depression, helplessness and self-deprecation. Prisoner participation in the social reality which the guards had structured for them lent increasing validity to it and, as the prisoners became resigned to their treatment over time, many acted in ways to justify their fate at the hands of the guards, adopting attitudes and behaviour which helped to sanction their victimisation. Most dramatic and distressing to us was the observation of the ease with which sadistic behaviour could be elicited in individuals who were not "sadistic types" and the frequency with which acute emotional breakdowns could occur in men selected precisely for their emotional stability.

Situational v. dispositional attribution

To what can we attribute these deviant behaviour patterns? If these reactions had been observed within the confines of an existing penal institution, it is probable that a dispositional hypothesis would be invoked as an explanation. Some cruel guards might be singled out as sadistic or passive-aggressive personality types who chose to work in a correctional institution because of the outlets provided for sanctioned aggression. Aberrant reactions on the part of the inmate population would likewise be viewed as an extrapolation from the prior social histories of these men as violent, anti-social, psychopathic, unstable character types.

Existing penal institutions may be viewed as *natural experiments* in social control in which any attempts at providing a causal attribution for observed behaviour hopelessly confound dispositional and situational causes. In contrast, the design of our study minimised the utility of trait or prior social history explanations by means of judicious subject selection and random assignment to roles. Considerable effort and care went into determining the composition of the

final subject population from which our guards and prisoners were drawn. Through case histories, personal interviews and a battery of personality tests, the subjects chosen to participate manifested no apparent abnormalities, anti-social tendencies or social backgrounds which were other than exemplary. On every one of the scores of the diagnostic tests each subject scored within the normal-average range. Our subjects then, were highly representative of middle-class, Caucasian American society (17 to 30 years in age), although above average in both intelligence and emotional stability.

Nevertheless, in less than one week their *behaviour* in this simulated prison could be characterised as pathological and anti-social. The negative, anti-social reactions observed were not the product of an environment created by combining a collection of deviant personalities, but rather, the result of an intrinsically pathological situation which could distort and rechannel the behaviour of essentially normal individuals. The abnormality here resided in the psychological nature of the situation and not in those who passed through it. Thus, we offer another instance in support of Mischel's [4] social-learning analysis of the power of situational variables to shape complex social behaviour. Our results are also congruent with those of Milgram [5] who most convincingly demonstrated the proposition that evil acts are not necessarily the deeds of evil men, but may be attributable to the operation of powerful social forces. Our findings go one step further, however, in removing the immediate presence of the dominant experimenter-authority figure, giving the subjects-as-guards a freer range of behavioural alternatives, and involving the participants for a much more extended period of time.

Despite the evidence favouring a situational causal analysis in this experiment, it should be clear that the research design actually *minimised* the effects of individual differences by use of a homogenous middle-range subject population. It did not allow the strongest possible test of the relative utility of the two types of explanation. We cannot say that personality differences do not have an important effect on behaviour in situations such as the one reported here. Rather, we may assert that the variance in behaviour observed could be reliably attributed to variations in situational rather than personality variables. The inherently pathological characteristics of the prison situation itself, at least as functionally simulated in our study, were a *sufficient* condition to produce aberrant, anti-social behaviour. (An alternative design which would maximise the potential operation of personality or dispositional variables would assign subjects who were extreme on pre-selected personality dimensions to each of the two experimental treatments. Such a design would, however, require a larger subject population and more resources than we had available.)

The failure of personality assessment variables to reliably discriminate the various patterns of prison behaviour, guard reactions as well as prisoner coping styles is reminiscent of the inability of personality tests to contribute to an understanding of the psychological differences between American P.O.W.s in Korea who succumbed to alleged Chinese Communist brain-washing by

"collaborating with the enemy" and those who resisted [6]. It seems to us that there is little reason to expect paper-and-pencil behavioural reactions on personality tests taken under "normal" conditions to generalise into coping behaviours under novel, stressful or abnormal environmental conditions. It may be that the best predictor of behaviour in situations of stress and power, as occurs in prisons, is overt behaviour in functionally comparable simulated environments.

In the situation of imprisonment faced by our subjects, despite the potent situational control, individual differences were nevertheless manifested both in coping styles among the prisoners and in the extent and type of aggression and exercise of power among the guards. Personality variables, conceived as learned behaviour styles can act as moderator variables in allaying or intensifying the impact of social situational variables. Their predictive utility depends upon acknowledging the inter-active relationship of such learned dispositional tendencies with the eliciting force of the situational variables.

Reality of the simulation

At this point it seems necessary to confront the critical question of "reality" in the simulated prison environment: were the behaviours observed more than the mere acting out assigned roles convincingly? To be sure, ethical, legal and practical considerations set limits upon the degree to which this situation could approach the conditions existing in actual prisons and penitentiaries. Necessarily absent were some of the most salient aspects of prison life reported by criminologists and documented in the writing of prisoners [7, 8]. There was no involuntary homosexuality, no racism, no physical beatings, no threat to life by prisoners against each other or the guards. Moreover, the maximum anticipated "sentence" was only two weeks and, unlike some prison systems, could not be extended indefinitely for infractions of the internal operating rules of the prison.

In one sense, the profound psychological effects we observed under the relatively minimal prison-like conditions which existed in our mock prison make the results even more significant and force us to wonder about the devastating impact of chronic incarceration in real prisons. Nevertheless, we must contend with the criticism that the conditions which prevailed in the mock prison were too minimal to provide a meaningful analogue to existing prisons. It is necessary to demonstrate that the participants in this experiment transcended the conscious limits of their preconceived stereotyped roles and their awareness of the artificiality and limited duration of imprisonment. We feel there is abundant evidence that virtually all of the subjects at one time or another experienced reactions which went well beyond the surface demands of role-playing and penetrated the deep structure of the psychology of imprisonment.

Although instructions about how to behave in the roles of guard or prisoner were not explicitly defined, demand characteristics in the experiment obviously exerted some directing influence. Therefore, it is enlightening to look to

circumstances where role demands were minimal, where the subjects believed they were not being observed, or where they should not have been behaving under the constraints imposed by their roles (as in "private" situations), in order to assess whether the role behaviours reflected anything more than public conformity or good acting.

When the private conversations of the prisoners were monitored, we learned that almost all (a full 90%) of what they talked about was directly related to immediate prison conditions, that is, food, privileges, punishment, guard harassment, etc. Only one-tenth of the time did their conversations deal with their life outside the prison. Consequently, although they had lived together under such intense conditions, the prisoners knew surprisingly little about each other's past history or future plans. This excessive concentration on the vicissitudes of their current situation helped to make the prison experience more oppressive for the prisoners because, instead of escaping from it when they had a chance to do so in the privacy of their cells, the prisoners continued to allow it to dominate their thoughts and social relations. The guards too, rarely exchanged personal information during their relaxation breaks. They either talked about "problem prisoners", or other prison topics, or did not talk at all. There were few instances of any personal communication across the two role groups. Moreover, when prisoners referred to other prisoners during interviews, they typically deprecated each other, seemingly adopting the guards' negative attitude.

From post-experimental data, we discovered that when individual guards were alone with solitary prisoners and out of range of any recording equipment, as on the way to or in the toilet, harassment often was greater than it was on the "Yard". Similarly, video-taped analyses of total guard aggression showed a daily escalation even after most prisoners had ceased resisting and prisoner deterioration had become visibly obvious to them. Thus guard aggression was no longer elicited as it was initially in response to perceived threats, but was emitted simply as a "natural" consequence of being in the uniform of a "guard" and asserting the power inherent in that role. In specific instances we noted cases of a guard (who did not know he was being observed) in the early morning hours pacing the "Yard" as the prisoners slept—vigorously pounding his night stick into his hand while he "kept watch" over his captives. Or another guard who detained an "incorrigible" prisoner in solitary confinement beyond the duration set by the guards' own rules and then he conspired to keep him in the hole all night while attempting to conceal this information from the experimenters who were thought to be too soft on the prisoners.

In passing, we may note an additional point about the nature of role-playing and the extent to which actual behaviour is "explained away" by reference to it. It will be recalled that many guards continued to intensify their harassment and aggressive behaviour even after the second day of the study, when prisoner deterioration became marked and visible and emotional breakdowns began to occur (in the presence of the guards). When questioned after the study about their persistent affrontive and harrassing behaviour in the face of prisoner

emotional trauma, most guards replied that they were "just playing the role" of a tough guard, although none ever doubted the magnitude or validity of the prisoners' emotional response. The reader may wish to consider to what extremes an individual may go, how great must be the consequences of his behaviour for others, before he can no longer rightfully attribute his actions to "playing a role" and thereby abdicate responsibility.

When introduced to a Catholic priest, many of the role-playing prisoners referred to themselves by their prison number rather than their Christian names. Some even asked him to get a lawyer to help them get out. When a public defender was summoned to interview those prisoners who had not yet been released, almost all of them strenuously demanded that he "bail" them out immediately.

One of the most remarkable incidents of the study occurred during a parole board hearing when each of five prisoners eligible for parole was asked by the senior author whether he would be willing to forfeit all the money earned as a prisoner if he were to be paroled (released from the study). Three of the five prisoners said, "yes", they would be willing to do this. Notice that the original incentive for participating in the study had been the promise of money, and they were, after only four days, prepared to give this up completely. And, more surprisingly, when told that this possibility would have to be discussed with the members of the staff before a decision could be made, each prisoner got up quietly and was escorted by a guard back to his cell. If they regarded themselves simply as "subjects" participating in an experiment for money, there was no longer any incentive to remain in the study and they could have easily escaped this situation which had so clearly become aversive for them by quitting. Yet, so powerful was the control which the situation had come to have over them, so much a reality had this simulated environment become, that they were unable to see that their original and singular motive for remaining no longer obtained, and they returned to their cells to await a "parole" decision by their captors.

The reality of the prison was also attested to by our prison consultant who had spent over 16 years in prison, as well as the priest who had been a prison chaplain and the public defender who were all brought into direct contact with our simulated prison environment. Further, the depressed affect of the prisoners, the guards' willingness to work overtime for no additional pay, the spontaneous use of prison titles and I.D. numbers in non role-related situations all point to a level of reality as real as any other in the lives of all those who shared this experience.

To understand how an illusion of imprisonment could have become so real, we need now to consider the uses of power by the guards as well as the effects of such power in shaping the prisoner mentality.

Pathology of power

Being a guard carried with it social status within the prison, a group identity (when wearing the uniform), and above all, the freedom to exercise an unprecedented degree of control over the lives of other human beings. This

control was invariably expressed in terms of sanctions, punishment, demands and with the threat of manifest physical power. There was no need for the guards to rationally justify a request as they do in their ordinary life and merely to make a demand was sufficient to have it carried out. Many of the guards showed in their behaviour and revealed in post-experimental statements that this sense of power was exhilarating.

The use of power was self-aggrandising and self-perpetuating. The guard power, derived initially from an arbitrary label, was intensified whenever there was any perceived threat by the prisoners and this new level subsequently became the baseline from which further hostility and harassment would begin. The most hostile guards on each shift moved spontaneously into the leadership roles of giving orders and deciding on punishments. They became role models whose behaviour was emulated by other members of the shift. Despite minimal contact between the three separate guard shifts and nearly 16 hours a day spent away from the prison, the absolute level of aggression as well as more subtle and "creative" forms of aggression manifested, increased in a spiralling function. Not to be tough and arrogant was to be seen as a sign of weakness by the guards and even those "good" guards who did not get as drawn into the power syndrome as the others respected the implicit norm of *never* contradicting or even interfering with an action of a more hostile guard on their shift.

After the first day of the study, practically all prisoner's rights (even such things as the time and conditions of sleeping and eating) came to be redefined by the guards as "privileges" which were to be earned for obedient behaviour. Constructive activities such as watching movies or reading (previously planned and suggested by the experimenters) were arbitrarily cancelled until further notice by the guards—and were subsequently never allowed. "Reward", then became granting approval for prisoners to eat, sleep, go to the toilet, talk, smoke a cigarette, wear glasses or the temporary diminution of harassment. One wonders about the conceptual nature of "positive" reinforcement when subjects are in such conditions of deprivation, and the extent to which even minimally acceptable conditions become rewarding when experienced in the context of such an impoverished environment.

We might also question whether there are meaningful non-violent alternatives as models for behaviour modification in real prisons. In a world where men are either powerful or powerless, everyone learns to despise the lack of power in others and in oneself. It seems to us, that prisoners learn to admire power for its own sake—power becoming the ultimate reward. Real prisoners soon learn the means to gain power whether through ingratiation, informing, sexual control of other prisoners or development of powerful cliques. When they are released from prison, it is unlikely they will ever want to feel so powerless again and will take action to establish and assert a sense of power.

The pathological prisoner syndrome

Various coping strategies were employed by our prisoners as they began to react to their perceived loss of personal identity and the arbitrary control of their

lives. At first they exhibited disbelief at the total invasion of their privacy, constant surveillance and atmosphere of oppression in which they were living. Their next response was rebellion, first by the use of direct force, and later with subtle divisive tactics designed to foster distrust among the prisoners. They then tried to work within the system by setting up an elected grievance committee. When that collective action failed to produce meaningful changes in their existence, individual self-interests emerged. The breakdown in prisoner cohesion was the start of social disintegration which gave rise not only to feelings of isolation but depreciation of other prisoners as well. As noted before, half the prisoners coped with the prison situation by becoming extremely disturbed emotionally—as a passive way of demanding attention and help. Others became excessively obedient in trying to be "good" prisoners. They sided with the guards against a solitary fellow prisoner who coped with his situation by refusing to eat. Instead of supporting this final and major act of rebellion, the prisoners treated him as a trouble-maker who deserved to be punished for his disobedience. It is likely that the negative self-regard among the prisoners noted by the end of the study was the product of their coming to believe that the continued hostility toward all of them was justified because they "deserved it" [9]. As the days wore on, the model prisoner reaction was one of passivity, dependence and flattened affect.

Let us briefly consider some of the relevant processes involved in bringing about these reactions.

Loss of personal identity. Identity is, for most people, conferred by social recognition of one's uniqueness, and established through one's name, dress, appearance, behaviour style and history. Living among strangers who do not know your name or history (who refer to you only by number), dressed in a uniform exactly like all other prisoners, not wanting to call attention to one's self because of the unpredictable consequences it might provoke—all led to a weakening of self identity among the prisoners. As they began to lose initiative and emotional responsivity, while acting ever more compliantly, indeed, the prisoners became deindividuated not only to the guards and the observers, but also to themselves.

Arbitrary control. On post-experimental questionnaires, the most frequently mentioned aversive aspect of the prison experience was that of being subjugated to the apparently arbitrary, capricious decisions and rules of the guards. A question by a prisoner as often elicited derogation and aggression as it did a rational answer. Smiling at a joke could be punished in the same way that failing to smile might be. An individual acting in defiance of the rules could bring punishment to innocent cell partners (who became, in effect, "mutually yoked controls"), to himself, or to all.

As the environment became more unpredictable, and previously learned assumptions about a just and orderly world were no longer functional, prisoners ceased to initiate any action. They moved about on orders and when in their cells rarely engaged in any purposeful activity. Their zombie-like reaction was the functional equivalent of the learned helplessness phenomenon reported by

Seligman and Groves [10]. Since their behaviour did not seem to have any contingent relationship to environmental consequences, the prisoners essentially gave up and stopped behaving. Thus the subjective magnitude of aversiveness was manipulated by the guards not in terms of physical punishment but rather by controlling the psychological dimension of environmental predictability [11].

Dependency and emasculation. The network of dependency relations established by the guards not only promoted helplessness in the prisoners but served to emasculate them as well. The arbitrary control by the guards put the prisoners at their mercy for even the daily, commonplace functions like going to the toilet. To do so, required publicly obtained permission (not always granted) and then a personal escort to the toilet while blindfolded and handcuffed. The same was true for many other activities ordinarily practised spontaneously without thought, such as lighting up a cigarette, reading a novel, writing a letter, drinking a glass of water or brushing one's teeth. These were all privileged activities requiring permission and necessitating a prior show of good behaviour. These low level dependencies engendered a regressive orientation in the prisoners. Their dependency was defined in terms of the extent of the domain of control over all aspects of their lives which they allowed other individuals (the guards and prison staff) to exercise.

As in real prisons, the assertive, independent, aggressive nature of male prisoners posed a threat which was overcome by a variety of tactics. The prisoner uniforms resembled smocks or dresses, which made them look silly and enabled the guards to refer to them as "sissies" or "girls". Wearing these uniforms without any underclothes forced the prisoners to move and sit in unfamiliar, feminine postures. Any sign of individual rebellion was labelled as indicative of "incorrigibility" and resulted in loss of privileges, solitary confinement, humiliation or punishment of cell mates. Physically smaller guards were able to induce stronger prisoners to act foolishly and obediently. Prisoners were encouraged to belittle each other publicly during the counts. These and other tactics all served to engender in the prisoners a lessened sense of their masculinity (as defined by their external culture). It follows then, that although the prisoners usually outnumbered the guards during line-ups and counts (nine v. three) there never was an attempt to directly overpower them. (Interestingly, after the study was terminated, the prisoners expressed the belief that the basis for assignment to guard and prisoner groups was physical size. They perceived the guards were "bigger", when, in fact, there was no difference in average height or weight between these randomly determined groups.)

In conclusion, we believe this demonstration reveals new dimensions in the social psychology of imprisonment worth pursuing in future research. In addition, this research provides a paradigm and information base for studying alternatives to existing guard training, as well as for questioning the basic operating principles on which penal institutions rest. If our mock prison could generate the extent of pathology it did in such a short time, then the punishment of being imprisoned in a real prison does not "fit the crime" for

most prisoners—indeed, it far exceeds it! Moreover, since prisoners and guards are locked into a dynamic, symbiotic relationship which is destructive to their human nature, guards are also society's prisoners.

Shortly after our study was terminated, the indiscriminate killings at San Quentin and Attica occurred, emphasising the urgency for prison reforms that recognise the dignity and humanity of both prisoners and guards who are constantly forced into one of the most intimate and potentially deadly encounters known to man.

Acknowledgments

This research was funded by an ONR grant: N00014-67-A-0112-0041 to Professor Philip G. Zimbardo.

The ideas expressed in this paper are those of the authors and do not imply endorsement of ONR or any sponsoring agency. We wish to extend our thanks and appreciation for the contributions to this research by David Jaffe who served as "warden" and pre-tested some of the variables in the mock prison situation. In addition, Greg White provided invaluable assistance during the data reduction phase of this study. Many others (most notably Carolyn Burkhart, Susie Phillips and Kathy Rosenfeld), helped at various stages of the experiment, with the construction of the prison, prisoner arrest, interviewing, testing, and data analysis—we extend our sincere thanks to each of these collaborators. Finally, we wish especially to thank Carlo Prescott, our prison consultant, whose personal experience gave us invaluable insights into the nature of imprisonment.

References

- 1 T. W. Adorno, E. Frenkel-Brunswick, D. J. Levinson & R. N. Sanford. *The Authoritarian Personality*. New York, Harper. 1950.
- 2 R. Christie & F. L. Geis (Eds). *Studies in Machiavellianism*. New York, Academic Press. 1970.
- 3 A. L Comrey. *Comrey Personality Scales*. San Diego, Educational and Industrial Testing Service. 1970.
- 4 W. Mischel. *Personality and Assessment*. New York, Wiley. 1968.
- 5 S. Milgram. Some conditions of obedience and disobedience to authority, *Human Relations* 1965, 18(1), 57-76.
- 6 G. Jackson. *Soledad Brother: the Prison Letters of George Jackson*. New York, Bantam Books. 1970.
- 7 E. Schein. *Coercive Persuasion*. New York, Norton. 1961.
- 8 H. Charrière. *Papillion*. Paris, Robert Laffont. 1969.
- 9 E. Walster. Assignment of responsibility for an accident, *Journal of Personality and Social Psychology* 1966, 3(1), 73-79.
- 10 M. E. Seligman & D. P. Groves. Nontransient learned helplessness, *Psychonomic Science* 1970, 19(3), 191-192.
- 11 D. C. Glass & J. E. Singer. Behavioural after effects of unpredictable and uncontrollable aversive events, *American Scientist* 1972, 6(4), 457-465.

A Re-analysis of the Reliability of Psychiatric Diagnosis

BJPsych

The British Journal of Psychiatry

A Re-analysis of the Reliability of Psychiatric Diagnosis

ROBERT L. SPITZER and JOSEPH L. FLEISS

The British Journal of Psychiatry 1974 125: 341-347
Access the most recent version at doi:[10.1192/bj.125.4.341](https://doi.org/10.1192/bj.125.4.341)

References

Article cited in:
<http://bjp.rcpsych.org#otherarticles>

Reprints/ permissions

To obtain reprints or permission to reproduce material from this paper, please write to permissions@rcpsych.ac.uk

You can respond to this article at

<http://bjp.rcpsych.org/cgi/eletter-submit/125/587/341>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

Downloaded from

bjp.rcpsych.org on February 9, 2011
Published by [The Royal College of Psychiatrists](#)

A Re-analysis of the Reliability of Psychiatric Diagnosis

By ROBERT L. SPITZER and JOSEPH L. FLEISS

INTRODUCTION

Classification systems such as diagnosis have two primary properties, reliability and validity. Reliability refers to the consistency with which subjects are classified; validity, to the utility of the system for its various purposes. In the case of psychiatric diagnosis, the purposes of the classification system are communication about clinical features, aetiology, course of illness and treatment. A necessary constraint on the validity of a system is its reliability. There is no guarantee that a reliable system is valid, but assuredly an unreliable system must be invalid.

Studies of the reliability of psychiatric diagnosis provide information on the upper limits of its validity. This paper discusses some of the difficulties in appraising diagnostic reliability, offers a re-analysis of available data from the literature, and suggests a possible course of action to improve psychiatric diagnosis.

Zubin (1967) reviewed the major studies of the reliability of psychiatric diagnosis performed until 1966. He noted that diagnostic reliability is referred to in three different ways: agreement between independent diagnosticians examining the same patients, stability in diagnosis over time, and similarity in diagnostic frequencies for comparable samples. It is the first sense—interjudge agreement—that is fundamental.

There are inherent limitations to the interpretation of the other two uses of the term. For agreement between initial and subsequent diagnosis, one must consider the possibility that some of the disagreement may be due to changes in the patient's condition and not just to unreliability. The difficulty with interpreting differences in distributions between populations is that one is forced to assume, often without evidence, that the populations do not differ in psychopathology, when in fact they may.

MEASURING DIAGNOSTIC RELIABILITY

More studies of diagnostic reliability have been of the interjudge type than of either of the other two types. There are two features of the data reported in these studies, however, which limit an assimilation of their results. One is the choice of an index of agreement and the other is a failure to take into account the base rates of the various diagnoses. The hypothetical data of Table I will illustrate some of the complexities involved in measuring diagnostic agreement.

TABLE I
Hypothetical data (in proportions) for agreement on three diagnoses by two diagnosticians

Diagnostician	Diagnostician B			Total
	A	Psychosis	Neurosis	
Psychosis ..	.75	.01	.04	.80
Neurosis ..	.05	.04	.01	.10
Organic ..	0	0	.10	.10
Total ..	.80	.05	.15	1.00

To measure the degree of agreement on a single diagnosis (e.g. neurosis), one may collapse the original table into a 2×2 table as shown in Table II. Some studies (Schmidt and Fonda, 1956; Kretzman, 1961) report the proportion of overall agreement, i.e., the proportion of all patients on whom there is agreement as to the

TABLE II
Hypothetical data (in proportions) for agreement on neurosis by two diagnosticians, from Table I

Diagnostician	Diagnostician B			Total
	A	Neurosis	Other	
Neurosis ..	.04	.06	.10	
Other ..	.01	.89	.90	
Total ..	.05	.95	1.00	

presence or the absence of the diagnosis. For the data of Table II, the proportion of overall agreement is $.04 + .89 = .93$.

Other studies (Beck *et al.*, 1962; Sandifer *et al.*, 1964) report the proportion of specific agreement, which is an index obtained by ignoring all subjects agreed upon as not having the given diagnosis (in Table II, ignoring the 89 per cent of patients agreed upon as not having a neurosis). One first determines the average proportion of all subjects given the specified diagnosis by either diagnostician (for the data of Table II, this proportion is $\frac{1}{2} (.10 + .05) = .075$), and then finds the proportion agreed upon as having that diagnosis (for the present example, this proportion is $.04$). The proportion of specific agreement is the ratio of these two proportions. For the data of Table II the resulting value is $.04/.075 = .53$. This index can be interpreted as the probability that one diagnostician will make the specified diagnosis given that the other has done so.

Table III presents the values of the two indices of agreement for the three diagnoses of Table I. The two indices order the diagnoses quite differently. The proportions of overall agreement seem to be similar, with that for organic brain syndrome being best and that for neurosis being second best. The proportions of specific agreement are of different orders of magnitude, and indicate that agreement on psychosis is best and agreement on neurosis poorest.

TABLE III
Indices of agreement between two diagnosticians on three diagnoses of Table I

Diagnosis	Proportion of overall agreement		Proportion of specific agreement		Kappa
	Chance	Obtained	Chance	Obtained	
	ex- pected	ex- pected	ex- pected	ex- pected	
Psychosis ..	.90	.68	.94	.80	.69
Neurosis ..	.93	.86	.53	.07	.50
Organic ..	.95	.78	.80	.12	.77

The two indices are obviously not comparable. A further complication is that neither can be interpreted independently of the rates at which the diagnoses are made. For one

thing, the values associated with the poorest possible agreement may be appreciably greater than 0. For example, given that the two diagnosticians diagnose psychosis 80 per cent of the time, the lowest value possible for the proportion of overall agreement on psychosis is .60 and the lowest value possible for the proportion of specific agreement on psychosis is .75.

Secondly, some degree of agreement is to be expected solely on the basis of chance. To take an extreme example, suppose that diagnosticians A and B jointly diagnosed a sample of patients without even examining them, but merely kept to their usual base rates. One would then expect that 64 per cent of the time ($= .8 \times .8$), they would agree on the diagnosis of psychosis. Given their base rates, only agreement beyond that expected by chance alone would be meaningful.

A statistic for measuring agreement on nominal categories such as diagnosis, which incorporates a correction for chance, was originally proposed by Cohen (1960) and later generalized by Spitzer *et al.* (1967a), Cohen (1968), Fleiss (1971), Light (1971), and Fleiss *et al.* (1972). The statistic, named kappa, contrasts the observed proportion of agreement with the proportion expected by chance alone by means of the formula $\kappa = (p_o - p_e) / (1 - p_e)$, where p_o is the observed proportion of agreement and p_e is the proportion expected by chance.

Whether p_o is taken to be the proportion of overall agreement or the proportion of specific agreement, one obtains identically equal values of kappa after correcting for chance. The term p_e is obtained by determining expected cell frequencies (as one does, e.g., in calculating the standard chi square statistic), and then calculating the proportion of agreement on the table with expected frequencies. Kappa varies from negative values for less than chance agreement, though 0 for chance agreement, to $+1.0$ for perfect agreement. Kappa may be interpreted as an intra-class correlation coefficient (Fleiss and Cohen, 1973).

Table III gives the chance expected values of the two proportions of agreement and the resulting values of kappa. The ordering effected by kappa is different from either of the other two order-

ings. After correcting for chance, one finds agreement to be best for organic brain syndrome, next best for psychosis, and poorest for neurosis.

STUDIES OF DIAGNOSTIC RELIABILITY

The major studies of the reliability of psychiatric diagnosis, fortunately, report both the base rates and the diagnostic agreement values, thus permitting the calculation of chance corrected agreement, kappa.

(I) Schmidt and Fonda (1956) studied 426 patients admitted to a state hospital in Connecticut. Each patient was diagnosed within the first week of admission by one of a group of eight psychiatric residents, and within the third week by one of three senior psychiatrists. The data available to the psychiatric residents were the usual admission reports as well as their own physical and mental status examination. The data available to the senior psychiatrist included all of the data available to the psychiatric residents as well as additional data that had been collected by other staff members and by themselves during their own brief examinations.

(II) Kreitman (1961) studied 90 consecutive new referrals to an out-patient clinic in England. Each patient was interviewed by one of three consulting psychiatrists, and completely independently by one of two research psychiatrists. The only sources of information to both sets of psychiatrists were the patient, a family member and a letter of referral.

(III) Beck *et al.* (1962) studied 153 patients randomly selected from new referrals to two out-patient services in Philadelphia. Each patient was randomly assigned to be interviewed by two of four experienced psychiatrists. Each psychiatrist conducted an independent interview and apparently had no source of information other than the patient himself.

(IV) Sandifer *et al.* (1964) studied 91 patients from three hospitals in North Carolina. A psychiatric resident presented material about each patient to a group of ten experienced psychiatrists. Following each presentation the patient was interviewed by one of the 10 diagnosticians. After jointly observing the

patient, each diagnostician made his own diagnosis.

(V) The U.S.-U.K. Diagnostic Project (Cooper *et al.*, 1972) conducted a series of studies comparing diagnostic practice in the United States and the United Kingdom. In one study, 250 consecutive admissions to a single New York State mental hospital and 250 consecutive admissions to a London area mental hospital were diagnosed by the hospital physician according to his usual practices, and independently by members of the project, who used a structured interview schedule. In a second study, 192 consecutive admissions to nine New York State mental hospitals and 174 consecutive admissions to nine London area mental hospitals were studied similarly. Most of the project members had received their psychiatric training in London. Because the results of the two studies within for each city were similar, only mean agreement values for the New York and the London samples are reported. The agreement measured is between the project's and the hospitals' psychiatrists.

(VI) Spitzer *et al.* (in preparation) studied 100 consecutive admissions to the Washington Heights Community Service of the New York State Psychiatric Institute. Each patient was diagnosed by one of 15 admitting residents within the first few days of admission. Each patient was also diagnosed up to three months after admission by one of two supervising psychiatrists after reviewing the case record prepared by the admitting resident. No attempt was made to prevent the admitting therapist from discussing his diagnostic formulation with the supervising psychiatrist. It is assumed that such discussions often took place, though not invariably.

RESULTS

Table IV presents the values of kappa calculated from the data presented in the original reports. Values are reported only for those categories for which original data were provided. Although the different studies used slightly different classification schemes (American Psychiatric Association, 1952 and 1968; H.M. Stationery Office, 1968), the results are reported for broad categories whose definitions

TABLE IV
Kappa coefficients of agreement on broad and specific diagnostic categories from six studies

Category	Study						Mean
	I	II	III	IV	New York	London	
Mental deficiency72		.72
Organic brain syndrome ..		.82	.90				.77
Acute brain syndrome ..				.44			.44
Chronic brain syndrome ..				.64			.64
Alcoholism74	.68	.71
Psychosis73	.62			
Schizophrenia77		.42	.32	.55
.57					.68	.60	
Affective disorder19	.44	.59
Neurotic depression47		.20	.26
Psychotic depression19	.24	.24
Manic-depressive33		.33
Involutional depression38	.21		.30
Personality disorder or Neurosis63			.51	.24	.44
Personality disorder33	.56	.19	.32
Sociopathic53			.53
Neurosis52			.40
Anxiety reaction42	.26	.45
Psychophysiological reaction ..					.38		.38

are similar in all of the classification systems used.

There are no diagnostic categories for which reliability is uniformly high. Reliability appears to be only satisfactory for three categories: mental deficiency, organic brain syndrome (but not its subtypes), and alcoholism. The level of reliability is no better than fair for psychosis and schizophrenia and is poor for the remaining categories. Using uncorrected agreement values, Zubin (1967) found agreement on the combined category of personality disorder and neurosis to be almost as high as for psychosis. It is clear from Table IV that after correction for chance, agreement on the combined category is poorer than on psychosis.

With the exception of the U.S.-U.K. study (number V) of the New York hospitals, all the studies summarized here involved diagnosticians of similar background and training. In addition,

special efforts were made in some of the studies to have the participant diagnosticians come to some agreement on diagnostic principles prior to the beginning of the study. One would have expected these features of similar background and prior consensus on principles to contribute to good reliability. One can only assume, therefore, that agreement between heterogeneous diagnosticians of different orientations and backgrounds, as they act in routine clinical settings, is even poorer than is indicated in Table IV. Further, there appears to have been no essential change in diagnostic reliability over time (the studies summarized in Table IV were arrayed in chronological order).

DISCUSSIONS AND CONCLUSIONS

In spite of the obvious unreliability of psychiatric diagnosis, there exists evidence for sensitivity to and agreement on the major psychiatric

problems experienced by a patient. Gurland *et al.* (1972), in a detailed analysis of data on the patients in the U.S.-U.K. diagnostic study, found that hospital psychiatrists were sensitive to patient's psychopathology. A number of patients in the New York sample were identified by the project psychiatrists as suffering from severe depression but not from any signs of schizophrenia. The hospital psychiatrists diagnosed most of these severe depressives as schizophrenic, but treated the majority of them with anti-depressant medication or with ECT. The hospital staffs obviously recognized the depression in their patients, when it was present, but failed to incorporate that recognition into their diagnoses.

As one of its studies of diagnostic practice, the U.S.-U.K. diagnostic project showed videotape recordings of a small number of psychiatric interviews to large numbers of American, British, and Canadian psychiatrists (Copeland *et al.*, 1971; Kendell *et al.*, 1971; Sharpe *et al.*, in press). Some of the interviews gave rise to strikingly large diagnostic differences between the three countries; in one case the percentage of psychiatrists diagnosing schizophrenia ranged from 2 per cent in the British Isles to 69 per cent in the United States, the proportion for Canadian psychiatrists being intermediate. In another study, Sandifer *et al.* (1968) reported substantial diagnostic differences between American, English and Scottish psychiatrists.

The participant psychiatrists in the videotape studies also judged the presence or absence of technically described psychiatric signs and symptoms, and made ratings on the Inpatient Multidimensional Psychiatric Scales (IMPS) of Lorr *et al.* (1962), a series of 89 rating scales defined in non-technical language. As Katz *et al.* (1969) found in an earlier study, the U.S.-U.K. study found poor agreement between psychiatrists in judging the presence or absence of symptoms described in technical terms. With respect to ratings on the IMPS, however, there were striking similarities in the psychiatrists' perceptions of psychopathology. Although American psychiatrists tended to rate the presence of more severe pathology than their British and Canadian colleagues, all psychi-

atrists were in excellent agreement as to the most serious and the least serious problem areas. In other words, mean profiles across the factors of the IMPS were at different mean levels, but were effectively parallel. This parallelism obtained for each of the tapes shown, even though the profile for each tape highlighted different aspects of psychopathology.

The reliability of psychiatric diagnosis as it has been practised since at least the late 1950's is not good. It is likely that the reasons for diagnostic unreliability are the same now as when Beck *et al.* (1962) studied them. They found that a significant amount of the variability among diagnosticians was due to differences in how they elicited and evaluated the necessary information, and that an even larger amount was due to inherent weakness and ambiguities in the nomenclature. Since that time there have been two major innovations which may provide solutions to these problems.

Several investigators have developed structured interview schedules which an interviewer uses in his examination of the patient (Spitzer *et al.*, 1967b and 1970; Wing *et al.*, 1967). These techniques provide for a standardized sequence of topics, and ensure that variability among clinicians in how they conduct their interviews and in what topics they cover is kept to a minimum. For rating the pathology observed, these schedules contain pre-coded items which explicitly define the behaviours to be rated rather than relying on technical terms which have different meanings to different clinicians.

With respect to improving the nomenclature, the St. Louis group (Feighner *et al.*, 1972) has offered a system limited to 16 diagnoses for which they believe strong validity evidence exists, and for which specified requirements are provided. Whereas in the standard system the clinician determines to which of the various diagnostic stereotypes his patient is closest, in the St. Louis system the clinician determines whether his patient satisfies explicit criteria. For example, for a diagnosis of the depressive form of primary affective disorder the three requirements are dysphoric mood, a psychiatric illness lasting at least one month with no other pre-existing psychiatric condition, and at least

five of the following eight symptoms: poor appetite or weight loss; sleep difficulty; loss of energy; agitation or retardation; loss of interest in usual activities or decrease in sexual drive; feelings of self-reproach or guilt; complaints of or actually diminished ability to think or concentrate; and thoughts of death or suicide.

A consequence of the St. Louis approach is the necessity for an 'undiagnosed psychiatric disorder' category for those patients who do not meet any of the criteria for the specified diagnoses. In actual use, this category is applied to 20-30 per cent of newly-admitted in-patients.

These two approaches, structuring the interview and specifying all diagnostic criteria, are being merged in a series of collaborative studies on the psychobiology of the depressive disorders sponsored by the N.I.M.H. Clinical Research Branch. We are confident that this merging will result not only in improved reliability but in improved validity which is, after all our ultimate goal.

ACKNOWLEDGEMENTS

This research was supported in parts by grants MH 23864 and MH 23964 from the National Institute of Mental Health. Professor Jacob Cohen of New York University made many valuable contributions to this report.

REFERENCES

- AMERICAN PSYCHIATRIC ASSOCIATION (1952) *Diagnostic and Statistical Manual of the Mental Disorders*.
- (1968) *Diagnostic and Statistical Manual of the Mental Disorders. 2nd Edition*.
- BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J. E., & ERBAUGH, J. K. (1962). Reliability of psychiatric diagnoses: 2. A study of consistency of clinical judgments and ratings. *Amer. J. Psychiat.*, **119**, 351-7.
- COHEN, J. (1960) A coefficient of agreement for nominal scales. *Educ. psychol. Measmt.*, **20**, 37-46.
- (1968) Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.*, **70**, 213-20.
- COOPER, J. E., KENDELL, R. E., GURLAND, B. J., SHARPE, L., COPELAND, J. R. M. & SIMON, R. (1972) *Psychiatric Diagnosis in New York and London*. (U.S.-U.K. Diagnostic Project.) London: Oxford University Press.
- COPELAND, J. R. M., COOPER, J. E., KENDELL, R. E., & GOURLAY, A. J. (1971) Differences in usage of diagnostic labels amongst psychiatrists in the British Isles. *Brit. J. Psychiat.*, **118**, 629-40.
- FEIGHNER, J. P., ROBINS, E., GUZE, S. B., WOODRUFF, R. A., WINOKUR, G. & MUÑOZ, R. (1972) Diagnostic criteria for use in psychiatric research. *Arch. gen. Psychiat.*, **26**, 57-63.
- FLEISS, J. L. (1971) Measuring nominal scale agreement among many raters. *Psychol. Bull.*, **76**, 378-82.
- & COHEN, J. (1973) The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ. psychol. Measmt.*, **33**, 613-19.
- SPITZER, R. L., ENDICOTT, J. & COHEN, J. (1972) Quantification of agreement in multiple psychiatric diagnosis. *Arch. gen. Psychiat.*, **26**, 168-71.
- GURLAND, B. J., FLEISS, J. L., SHARPE, L., SIMON, R. & BARRETT, J. E. (1972) The mislabeling of depressed patients in New York State hospitals. *Disorders of Mood* (eds. J. Zubin & F. A. Freyhan), pp. 17-28. Baltimore: Johns Hopkins Press.
- HER MAJESTY'S STATIONERY OFFICE (1968) *A Glossary of Mental Disorders*. General Register Office Studies on Medical and Population Subjects, no. 22.
- KATZ, M. M., COLE, J. O. & LOWRY, H. A. (1969) Studies of the diagnostic process: The influence of symptom perception, past experience, and ethnic background on diagnostic decisions. *Amer. J. Psychiat.*, **125**, 937-47.
- KENDELL, R. E., COOPER, J. E., GOURLAY, A. J., COPELAND, J. R. M., SHARPE, L. & GURLAND, B. J. (1971) The diagnostic criteria of American and British psychiatrists. *Arch. gen. Psychiat.*, **25**, 123-30.
- KREITMAN, N. (1961) The reliability of psychiatric diagnosis. *J. ment. Sci.*, **107**, 876-86.
- LIGHT, R. J. (1971) Measures of agreement for qualitative data: Some generalizations and alternatives. *Psychol. Bull.*, **76**, 365-77.
- LORR, M., McNAIR, D. M., KLETT, C. J. & LASKY, J. J. (1962) Evidence of ten psychiatric syndromes. *J. consult. Psychol.*, **26**, 185-9.
- SANDIFER, M. G., HORDERN, A., TIMBURY, G. C. & GREEN, L. M. (1968) Psychiatric diagnosis: A comparative study in North Carolina, London and Glasgow. *Brit. J. Psychiat.*, **114**, 1-9.
- PETTUS, C. & QUADE, D. (1964) A study of psychiatric diagnosis. *J. nerv. ment. Dis.*, **139**, 350-6.
- SCHMIDT, H. O. & FONDA, C. P. (1956) The reliability of psychiatric diagnosis: A new look. *J. abnor. soc. Psychol.*, **52**, 262-7.
- SHARPE, L., GURLAND, B. J., FLEISS, J. L., KENDELL, R. E., COOPER, J. E. & COPELAND, J. R. M. Some comparisons of American, Canadian and British psychiatrists in their diagnostic concepts. *Canad. J. Psychiat.* In press.
- SPITZER, R. L., COHEN, J., FLEISS, J. L. & ENDICOTT, J. (1967a) Quantification of agreement in psychiatric diagnosis: A new approach. *Arch. gen. Psychiat.*, **17**, 83-7.
- ENDICOTT, J., COHEN, J. & FLEISS, J. L. Constraints on the validity of computer diagnosis. (In preparation).
- — — FLEISS, J. L. & COHEN, J. (1970) Psychiatric Status Schedule: A technique for evaluating psychopathology and impairment in role functioning. *Arch. gen. Psychiat.*, **23**, 41-55.
- FLEISS, J. L., ENDICOTT, J. & COHEN, J. (1967b) Mental Status Schedule: Properties of factor analytically derived scales. *Arch. gen. Psychiat.*, **16**, 479-93.

- WING, J. K., BIRLEY, J. L. T., COOPER, J. E., GRAHAM, P. & ISAACS, A. D. (1967) Reliability of a procedure for measuring and classifying 'present psychiatric state'. *Brit. J. Psychiat.*, 113, 499-515.
- ZUBIN, J. (1967) Classification of the behavior disorders. In *Annual Review of Psychology* (eds. P. R. Farnsworth & O. McNemar). Palo Alto, California, *Annual Reviews*, pp. 373-406.

A synopsis of this paper was published in the June 1974 *Journal*.

Robert L. Spitzer, M.D., Director of Evaluation Section, Biometrics Research, New York State Department of Mental Hygiene at the Psychiatric Institute, 722 West 168 Street, New York, New York 10032; and Associate Professor of Clinical Psychiatry, Columbia University, New York

Joseph L. Fleiss, Ph.D., Head of Biostatistics Section, Biometrics Research, New York State Department of Mental Hygiene at the Psychiatric Institute; and Associate Professor of Biostatistics, Columbia University, New York

(Received 17 January 1974)

B *Lab Activities*

Accuracy of Observation

among
mple,
roducts

—for

ian, &
few

avior.

ay,

iversity

ences.

ACTIVITY

2

ACCURACY OF OBSERVATION

Paul J. Woods

Concept

This activity provides an intriguing introduction to discussions of the accuracy and reliability of humans as observers of behavioral events and environmental characteristics.

Materials Needed

A portable tape recorder, some imagination, and decent weather.

Instructions

Introduce the activity to the students simply by explaining that you are going to test their powers of observation—their skill as observers. Tell them that you are going to ask a colleague of yours (or a particular member of the class) to lead the class on a 15- or 20-minute journey around the school grounds. Inform the students that nothing unusual has been planned, and instruct them simply to observe the normal activities and circumstances in which they find themselves. They can be told that following the walk, you will be asking them some questions about their observations.

As your colleague or class member is leading the class on the journey around the campus, follow a short distance behind—close enough to be able to observe the class and its environment but sufficiently removed to be able to speak quietly into a portable tape recorder without being heard by the class. As the walk proceeds, record approximately 50–60 questions on the tape, spontaneously drawn from your own observations of events and objects. (You may wish to prearrange the route to be followed and walk it through once by yourself to practice the questioning routine.) Each question you record should be followed immediately by the answer. Questions should be completely factual and objectively confirmable. Following are some examples:

“How many planes flew overhead as we reached the playing field?”

“In what activity was the first group of students encountered by the class engaged?”

“Where was the class when the nearby truck backfired?”

“What color was the car that passed the class at the entrance to the parking lot?”

Avoid questions with answers that are a matter of opinion (“Was it a nice day?” “Did it seem warm?”).

Upon returning to the classroom, play each recorded question in turn and stop the recorder before the correct answer is played. Have students write their answers down on paper, and then play back the correct answer. On the chalkboard, keep a record of the number of right and wrong answers to each question.

Discussion

The broad variation in student responses should provide gist for a discussion of (a) the reliability of human observations, (b) the “truth” of courtroom testimony, (c) the benefits of structuring the collection of

Suggested Background Readings

observations (perhaps by specifying beforehand what information is to be sought), and (d) the value and necessity of independent replication and confirmation of answers. Comparing the accuracy of the teacher's answers with that of the students' answers should provide compelling demonstration of the value of specifying the information to be sought before it is collected.

- Bickman, L. Observational methods. In C. Sellitz, L. S. Wrightsman, & S. W. Cook, *Research methods in social relations* (3rd ed.). New York: Holt, Rinehart & Winston, 1976.
- Hutt, S. J. & Hutt, C. *Direct observation and measurement of behavior*. Springfield, Ill.: Charles C Thomas, 1970.
- Levin, M. *Understanding psychological research*. New York: Wiley, 1979. (chap. 11)
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally, 1966. (chaps. 5, 6)

Con

Mate

Instr

Operant Conditioning: Role in Human Behavior

ACTIVITY
23

OPERANT CONDITIONING: ROLE IN HUMAN BEHAVIOR

Edward Stork

Concept

From infancy onward, conditioning plays a major role in our lives. Yet most of us tend to downplay that role, possibly feeling that to admit such control over our behavior would be to admit that our lives are overly determined. Often when students read in their texts about classical and operant conditioning, they tend to associate that type of learning with infrahuman animals. That is, "Dogs, rats, and pigeons are affected by conditioning, but it doesn't play any role in my behavior." This activity is designed to provide a starting point for discussion of conditioning in humans.

Instructions

While discussing operant conditioning, interrupt your lecture with "Oh, by the way, before I forget again" and then ask a question to which you know you will get either an almost totally positive or negative response. For example, if your students are primarily seniors, you might say, "I was supposed to ask, how many of you have signed for a diploma for graduation?" All students will usually raise a hand. Then tell them to hold the position they are in and ask if anyone told them to raise their hands or even mentioned raising hands.

Discussion

The usual response is a chorus of groans as the students recognize that they have been "used." Discuss the activity as an example of human conditioning. Ask students to generate other examples that describe conditioning in humans. You can use the ensuing discussion as a bridge to talking about conditioning techniques used with humans in behavior therapy.

Suggested Background Readings

- Bellack, A. S., & Hersen, M. *Behavior modification: An introductory textbook*. New York: Oxford University Press, 1977.
Hulse, S. H., Deese, J. E., & Smith, H. E. *The psychology of learning* (5th ed.). New York: McGraw-Hill, 1980.
Smith, W. I. *Conditioning and instrumental learning* (2nd ed.). New York: McGraw-Hill, 1978.

Operant Conditioning Demonstration

ACTIVITY

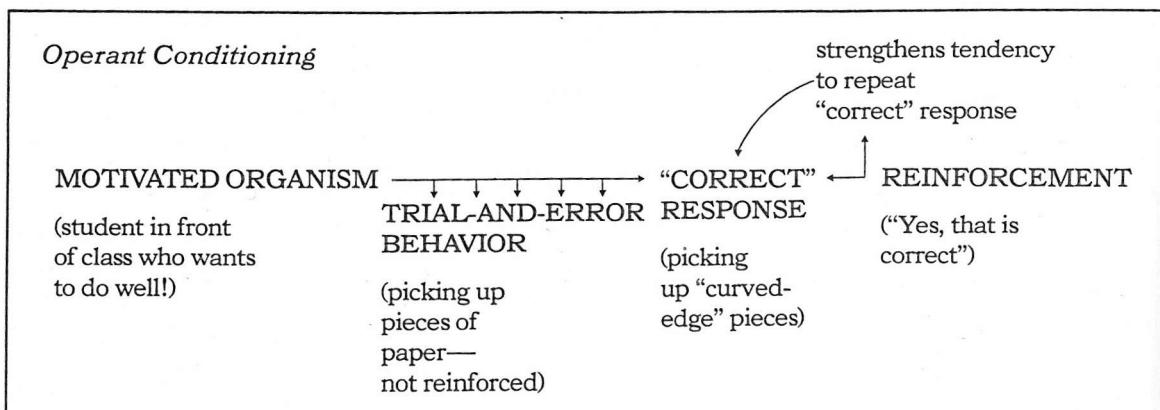
24

OPERANT CONDITIONING DEMONSTRATION

Patricia Keith-Spiegel

Concept

This classroom demonstration is a simple way to illustrate the process of operant (instrumental) conditioning with a human being. (It has never failed in 14 years, although some volunteers take longer than others to "learn.")

**Instructions**

Cut out approximately 40 pieces of paper 2 inches \times 2 inches. Then cut these 2-inch squares into a variety of shapes. Twenty of the pieces should have only straight edges, and 20 pieces should have at least one curved side each. (Save them in an envelope for repeated use.) Mix up the paper shapes and spread them out on a desk or table top. Ask for a student volunteer. Tell the volunteer to start picking up the pieces of paper one at a time and place them in a box. Each time the student picks one of the pieces that has a curve on it, say "Yes, that is correct." Each time the student picks up a completely straight-edged piece, give no reinforcement at all. Usually within 10 to 15 draws, the student will "learn" what has been defined as the "correct response" and will swiftly continue picking up pieces until all of the curved pieces are gone. (The students observing usually begin to approvingly giggle as the volunteer receives rapid positive reinforcements.)

Discussion

Ask the student to tell the class what she or he has learned. Be sure to point out to the class that the *only* input given the student (aside from the initial direction to pick up pieces) was in the form of Positive Reinforcement. Nothing was said about shapes or any other facet of the task. Ask the student to relate to the class what went through his or her mind during the learning process. Various trial-and-error strategies often emerge here (e.g., "At first I thought it was the larger pieces that were correct because the first big one I picked up was correct.") The basic operant-conditioning diagram provided here is helpful in summarizing for the class what they have just witnessed.

**Suggested
Background
Readings**

- Keller, F. S. *Learning: Reinforcement theory* (2nd ed.). New York: Random House, 1969.
- Krech, D., Crutchfield, R. S., Livson, N., & Krech, H. *Psychology: A basic course*. New York: Knopf, 1976. (chap. 3)
- Smith, W. I. *Conditioning and instrumental learning* (2nd ed.). New York: McGraw-Hill, 1978.

process of
is never
others to

y

EMENT

s. Then cut
ieces should
ne curved
ip the
for a
ieces of
udent picks
t." Each
give no
nt will
ill swiftly
one. (The
volunteer

sure to
le from the

cet of the
is or her
gies often
it were
basic
iarizing

Defense Mechanisms

ACTIVITY
75
DEFENSE MECHANISMS
Jack J. Greider

Concept

This activity is designed to increase student familiarity with the ways in which defense mechanisms are used. In addition to being a good teaching tool, the activity allows for the expression of a considerable amount of creativity on the part of the students. Defense mechanisms covered include regression, rationalization, repression, projection, fantasy, compensation, identification, and reaction formation.

Instructions

Ask for eight student volunteers for a role-playing exercise. I usually choose four males and four females, since I have found mixed pairs to work best in this activity. Have these eight students leave the room and go with them to explain what they are to do. First group them in male-female pairs and assign two of the defense mechanisms to each pair. Instruct each pair of students to make up two skits, one for each of the mechanisms. I usually allow them about 15 minutes to prepare their skits. While they are working, go back into the room, list the defense mechanisms on the chalkboard, and discuss each one briefly with the class. Then call the role players back into the room to put on their skits. Have each pair perform one of their two skits, and after going through the pairs have each pair do their second skit. The job of the class is to guess which mechanism each skit represents.

Discussion

In discussion you can point out the place of defense mechanisms in personality theory. You should comment on the role of defense mechanisms in normal functioning. When are these mechanisms helpful? When are they harmful? Conduct an anonymous survey of your students asking them to indicate whether or not they have used any of the defense mechanisms during the past year. Do they use some defense mechanisms more than others? If so, which ones?

**Suggested
Background
Readings**

- Clum, G. A., & Clum, J. Mood variability and defense mechanism preference. *Psychological Reports*, 1973, 32, 910.
Houston, J. P., Bee, H., Hatfield, E., & Rimm, D. C. *Invitation to psychology*. New York: Academic Press, 1979. (chap. 11)
Viney, L. L., & Manton, M. Defense mechanism preferences and the expression of anxiety. *Social Behavior and Personality*, 1974, 2, 50-55.
Weiner, B. et al. *Discovering psychology*. New York: St. Martin's Press, 1977. (chap. 15)

Concept**Instructions**

To Sleep, Perchance to Dream

ACTIVITY 80

TO SLEEP, PERCHANCE TO DREAM

Ludy T. Benjamin, Jr.

Concept

There are probably more myths and misunderstandings about sleep and dreaming than about any other aspect of human behavior except sex. For most people, sleep and dreaming will occupy one third of their lives, yet few individuals are aware of even the most rudimentary information about this activity. For example, consider the following facts: Apparently everyone dreams; that is, there are no nondreamers, only nonrecallers. In an 8-hour night of sleep most people will have from four to five dreams, usually on quite unrelated topics. The great majority of dreams are in color. Sleep learning apparently does not occur. Dreams are not always filled with easily interpretable psychological meanings, as popular literature would have one believe.

The study of sleep and dreaming is still in its infancy. Most of what we know about this area is the result of research occurring in the last 25 years and is principally due to the development of electrophysiological techniques (particularly the electroencephalograph, or EEG) and the discovery, in 1953, of rapid eye movements (REMs) during sleep.

The purposes of this exercise are (a) to help students focus attention on their own sleep and dreaming patterns, (b) to generate data for class discussion, (c) to introduce students to the concept of data collection in sleep and dreaming, (d) to illustrate appropriate statistical measures for summarizing the data, and (e) to aid students in understanding the interpretation of data in general and these data in particular.

Preparation of Class

Pass out the "Sleep and Dreaming Record" to each student, providing multiple copies—one for each day that records will be kept (14 consecutive nights provides a good sample). Tell the students why the data are being collected and indicate that participation is voluntary. Explain that they can conceal their identity by marking their records with a number (six or seven digits in length) that they make up on their own. This procedure allows them to identify their statistics in reference to others in class when the summary data are provided later. Note that it takes only about 5 minutes each day to complete the record.

Instructions

Ask all students to begin their records on the same day and to keep their record sheets at home. Request that during the time the records are being kept, discussion among students regarding their sleep and dreaming patterns be minimal to avoid unintentional influences on the data. Give explicit instructions on record keeping. (It is a good idea to have one "practice run" to insure that everyone understands the record-keeping system before continuing for 14 days.) The record provided is only a sample. Feel free to modify it as appropriate.

When the time period is completed and students have turned in all of their records, the tedious part begins for the teacher—the

Discussion

summarization of the data, for which a calculator is most helpful. Each student's records should be summarized separately, and data should be analyzed for the group as a whole. A summary sheet for the students should then be prepared, which lists everyone who participated (by number) and provides the group analyses. For example, Student # 107654 can examine that column number of the summary sheet to find a mean sleep time of 7.3 hours (over 14 days), with a range of 5.6–9.7 hours. The student can then compare these figures with those of other students and with those for the class as a whole, based on the group data.

Sleep and Dreaming Record

Student Number _____ Date _____

1. Total sleep time (in hours) _____. On the time line below, block out your sleep periods, including naps.

6:00 10:00 2:00 6:00 10:00 2:00 6:00
p.m. p.m. a.m. a.m. a.m. p.m. p.m.

2. Total number of awakenings during major sleep period. _____. (Do not count the final morning awakening.)
3. On the scale below, rate the quality of your night's sleep (in your opinion). Circle one of the numbers from plus four to minus four.

bad -4 -3 -2 -1 0 +1 +2 +3 +4 good

4. In your judgment, how many separate dreams can you recall at least a fragment of? _____
5. It is possible that you will recall some of your dreams better than others. Using percentages, estimate the amount of each dream recalled.

Dream 1 _____ Dream 2 _____ Dream 3 _____ Dream 4 _____

6. How many of these dreams could you relate to presleep experiences of the dream day? _____
7. Did you appear as a character in the dreams you recall? In how many? _____
8. How many of your dreams were in color? _____
9. Were there stimuli in your dreams of a nonvisual nature? Check the following if appropriate.

sound _____ taste _____ touch _____ smell _____

The data students generate will add considerable personal interest to the topic of sleep and dreaming. Further, the statistical treatments will help them understand how data are summarized and analyzed to make them more meaningful. It might be useful to save the summary statistics from classes for comparison with those of future classes.

Discussion

Optional

Although it is not necessary, teachers may wish to provide some statistical measures of relationship between some of the variables for which they have collected data by using the technique of correlation. For example, is there a relationship between the number of hours people sleep and the number of dreams they recall? Or is there a relationship between the subjective sleep quality rating and the number of awakenings one experiences during the night? A number of correlational analyses can be computed to answer these and other questions. To compute these values, a calculator is needed. Some calculators have a built-in correlation function; otherwise, the computational formula in Appendix A should be used. It is important to remember that correlation is a measure of the degree to which two variables are related and does not necessarily specify the *nature* of the relationship. That is, one *cannot* assume that if two variables are shown by correlation to be related, that the relationship is one of cause and effect.

Concept**Suggested
Background
Readings**

Dement, W. C. *Some must watch while some must sleep*. San Francisco:

Freeman, 1974.

Webb, W. B. *Sleep: The gentle tyrant*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Instruction

Mental Illness

es for
ation. For
eople
onship

relational
To
have a
ula in
orrelation
d does
re cannot
ted, that

Francisco:
Prentice-

ACTIVITY

81

MENTAL ILLNESS

James M. Gardner

Concept

We often make judgments about the behavior of other people, determining whether particular acts were or were not appropriate. Occasionally we may even decide, based on our observations, that someone is "mentally ill." This exercise offers a classroom opportunity to study such attributions.

Instructions

Ask for six volunteers to play roles in a skit. Take them out of the room and have them choose a role from six 3 × 5 inch cards bearing the following descriptions:

1. You are an escaped convict, previously convicted of murder.
2. You are a successful business executive whose spouse just announced the existence of a love affair, whose child is in the hospital, and whose car broke down this morning.
3. You are married with two children and unemployed; you desperately need a job and are on your way to an interview.
4. You are on your way to a sale.
5. You are lonely, have few friends, became bored watching TV, and are going somewhere just to have something to do.
6. You are waiting at the bus stop for the bus. Your role is to try to engage each of the waiting passengers in conversation so that the class can observe how they act. Some questions that you might ask are: Do you have the proper time? Does the bus usually run late?

While the actors study their roles, return to the room and instruct the other students to observe closely the behavior of each actor. Bring in the actor who drew the sixth role—the catalyst. Then bring in the other actors one by one, and let them interact with the catalyst for about 1 minute each.

When the skit is over, have the actors return to their seats. On the chalkboard, set up a matrix with the names of the actors in a vertical column on the left; list them in any order, except put the catalyst's name last. Tell the class that the six actors were playing roles, that you are going to list those roles across the top of the matrix, and that you are going to poll the class to determine how many believe each actor could have been playing each role. Also tell them that you are going to include a role that no one played. Then list the roles one at a time, polling students on each one before you list the next. List the roles in the same order as they appear above, but insert "a mentally ill person" between the third and fourth roles.

Adapted from the "Mental Illness Game" by James M. Gardner, *Teaching of Psychology*, 1976, 3(3), 141–142. Copyright 1976 by Division Two of the American Psychological Association. Reprinted by permission.

When the voting is completed, examine the data. First, determine which actors appeared to play several roles and which were identified with only one role. Next, determine which roles were clearly identified and which appeared to be played by many actors. Then ask the actors to state which role they played. It will become apparent that no one was assigned the role of "a mentally ill person." Now check how many believed each of the six actors to be playing that role. Compare the total number here to totals for other roles. Usually the total for "a mentally ill person" is one of the highest.

Discussion

This outcome can lead to a discussion of how easy it is for people to be labeled mentally ill when there is no psychological problem present. Ask the students to verbalize what behaviors they observed in the actors that they felt were indicative of mental illness. Is there any agreement on those behaviors? Were those behaviors exhibited by actors who were not thought to be mentally ill? You can add a number of topics to the discussion such as witch hunts and the legal definition of insanity. You can also discuss the concept of mental illness. Does the use of the term *illness* promote a medical model? What other terms are used and how are they different?

Suggested Background Readings

- Gardner, J. M. *Community psychology: The left hand of the magician*. New York: Plenum, 1980.
Sahakian, W. S. (Ed.). *Psychopathology today*. Itasca, Ill.: F. E. Peacock, 1970.
Szasz, T. S. The myth of mental illness. *American Psychologist*, 1960, 15, 113-118.
Szasz, T. S. *The manufacture of madness*. New York: Delta, 1970.

Concept

Instruction

Classical Conditioning

19 JAWS: DEMONSTRATING CLASSICAL CONDITIONING

Randolph A. Smith

Ouachita Baptist University

This activity is a good “opener” to get students to pay attention and talk. The materials are readily available, and once you have them, there is no advance preparation. The music from the movie *Jaws* could be replaced by music from any current thriller, with appropriate modifications in imagery. No prior knowledge of psychology is necessary, and the activity is suitable for classes of all sizes.

- CONCEPT** The concept of classical conditioning is often difficult for students to grasp if the instructor immediately begins to use the time-honored example of Pavlov and his dogs. The students may get lost in the maze of terms—US, UR, CR, CS—and fail to see any relevance between slobbering dogs and human learning. Thus I think that it is important to begin a presentation of classical conditioning with a real-life example that most students have experienced.
- MATERIALS NEEDED** You will need a tape recording of a few seconds of the shark-attack music from the movie *Jaws*.
- INSTRUCTIONS** I use this demonstration at the beginning of the section on conditioning in general psychology classes before I introduce the concepts. First I ask students to close their eyes and engage in mental imagery. I tell them to imagine that it is a hot summer day and that they are at the beach. The sun is scorching. They are getting hotter and hotter, can stand it no longer, run toward the ocean, and splash in the shallow water. Then they swim out to deeper water and enjoy cooling off after being in the sun. It is fairly easy to stretch this imagery process out for 2 or 3 minutes. Then, as unobtrusively as possible, I start the music from *Jaws*. From the expressions on their faces and the laughter, I can tell immediately that most students have seen *Jaws* and have been conditioned to associate the killer shark with the music.
- DISCUSSION** The key part of the demonstration is the discussion that follows. Without knowing the appropriate classical conditioning terms, students are able to explain in their own words what happened during the movie. They can tell you that the shark (technically a CS, but a US in this demonstration) yields a response of fear or disgust (the UR). They know that the music (CS) originally had no meaning but came to elicit a response (CR) during the movie because it signaled that the shark was about to appear and have a swimmer for lunch. At this point, they are prepared to hear your discussion on Pav-

lov's experiment, to understand it and the terminology used, and to see the application of classical conditioning to human behavior. As you go on through the classical conditioning information, you can point out that you actually demonstrated second-order conditioning to them, because their responses to a shark are learned through classical conditioning since they associate sharks with mutilation, blood, death, and so forth. All in all, the demonstration helps to bring classical conditioning and its principles to life in a highly relevant situation.

SUGGESTED
READING

- Cogan, D., & Cogan, R. (1984). Classical salivary conditioning: An easy demonstration. *Teaching of Psychology, 11*, 170–171.
- Gibb, G. D. (1983). Making classical conditioning understandable through a demonstration technique. *Teaching of Psychology, 10*, 112–113.

Human Operant Conditioning

20

HUMAN OPERANT CONDITIONING

John K. Bare
Carleton College

This activity does not require prior knowledge of psychology. There is little advance preparation, and it is suitable for classes of virtually any size. You will need to give some thought to how you will instruct the "experimenter(s)" in the procedures without informing the learner(s) during this in-class exercise.

CONCEPT

Human operant conditioning is often demonstrated by asking the subject to say nouns and then reinforcing the plural but not the singular form. Another technique is to ask the subject to say numbers and then reinforce those numbers divisible by 2. The shortcoming of these procedures is that the subject is limited by the instructions to two classes of responses, whereas the animal in the typical Skinner box can make many more different responses even if the space is limited. In addition, the procedures do not display shaping of the response by successive approximation, and as a consequence, there are no changes in the criteria for reinforcement that are present in shaping. Moreover, during extinction, one cannot watch the behavior revert to previously rewarded responses. These shortcomings can be minimized by doing the activity outlined here.

MATERIALS NEEDED

No materials are required except pencil and paper. Score sheets must be constructed by those doing the reinforcing.

INSTRUCTIONS

Twelve words are shown to the subject, and his or her instructions are simply to try to get as many points as possible. He or she is to say the word, and following the word the experimenter will say either "Point" or "No point." The experimenter should record which word was given and whether a reinforcement was given for each trial. The words to be used are *underfed*, *misread*, *understand*, *understudy*, *mistake*, *misread*, *understanding*, *mistaken*, *underpaid*, *understandingly*, *misinform*, and *mistakenly*.

The experimenter first reinforces any word that begins with "mis," by saying "Point." Once the subject has picked three words in succession that begin with the syllable "mis" and received three reinforcements in succession, the experimenter should change the criterion so that reinforcement is received only if the first two syllables "mistake . . ." are chosen. Again, after the subject has received three reinforcements in succession, the criterion is changed so that only the word with the three syllables "mistaken . . ." is reinforced. After three more successive reinforcements, only the word "mistakenly" is reinforced. After five correct responses in succession, the reinforcement is no longer given for any word, thus beginning the extinction process. The extinction session should last at least three minutes.

The students can work in pairs, with one as the subject and one as the experimenter, or one student can be used to demonstrate the phenomena. Ask the subject to pick his or her words by number, because that will facilitate recording the responses. In order to draw a cumulative response curve, it is necessary to have someone record how many correct responses occur in each 3-minute period.

DISCUSSION

Once the cumulative response curves for acquisition and extinction have been drawn, have students make observations on the rate of responding (indicated by the slope). You might ask them to try to specify what a reacquisition curve might look like or a curve showing spontaneous recovery. If a number of subjects are conditioned, then individual differences in rate will occur, so you may wish to ask what is it about the behavior in the situation that might account for such individual differences. The regression back through previously reinforced responses that is often observed in animals may also occur. Ask the class how such behavior—going back through previously reinforced responses—might be adaptive for an animal in its natural environment.

**SUGGESTED
READING**

Gleitman, H. (1986). *Psychology* (2nd ed., pp. 101–113). New York: Norton. Nearly any introductory text will discuss operant or instrumental conditioning. See Gleitman's small section on shaping in pp. 107–108.

Bringing the Clinic into the Undergraduate Classroom

45

BRINGING THE CLINIC INTO THE UNDERGRADUATE CLASSROOM

David M. Young

Indiana University—Purdue University at Fort Wayne

Students develop hands-on skills in problem formulation, classification of disorders, developing treatment plans, and assessing the prognoses of actual or simulated clients in this exercise. Although a film or video presentation of a clinical interview works best, you can do this activity with an audiotaped interview or even a written history. The simulation procedure generally should be introduced 4 to 5 weeks into the course, after students have been prepared in the background they will need to engage in this miniclinic assessment exercise. A full class session for the actual presentation of the interview, the assessment, and ensuing discussion is recommended. The student worksheets for completing the assessment are provided.

CONCEPT Many notable teachers of psychology, textbook authors, and practicing psychologists (Benjamin & Lowman, 1981; McKeachie, 1978; Radford & Rose, 1980) have implored professors to give undergraduate students the experience of actually working at the tasks that psychologists perform. In many undergraduate lecture courses, students become oriented to what clinicians do and may even observe audiotaped or videotaped examples of clinical activities. Yet, because of ethical concerns (e.g., confidentiality problems, using untrained students to make interventions or important decisions regarding a client) or the sheer numbers of students involved, psychology students are rarely directed to attempt the hands-on exercise of clinical skills. When students are tested in abnormal psychology courses, they are often expected to make diagnostic decisions, provide a prognosis, and define the relevance of particular forms of therapy for different disorders—all based on fragments of hypothetical cases, often presented in the form of multiple-choice items. This article describes a method that involves students in the relevant simulation of a variety of clinical experiences. We have employed these techniques or "miniclinics" in several courses that contain units related to clinical assessment and intervention (e.g. abnormal, child development, and introductory psychology courses).

**MATERIALS
NEEDED**

Besides the three student worksheets that are replicated here, the most important element of the miniclinic exercise is the case material presented by the instructor—ideally, written background information and films or videotapes. Satisfactory but less revealing and stimulating for discussion are written case histories.

Recently, there has been an increase in both the quality and the selection of clinical vignettes (both actual and simulated productions). These are often available from book publishers for a small fee or at no charge to instructors who have adopted the publisher's text book in abnormal psychology.

The film catalogues of colleges and universities are another excellent source of such material. The film "Otto—A Case Study in Abnormal Behavior" (Film No. EC1404, 16; available from the Audio-Visual Center, Indiana University, Bloomington, IN 47405-5901) features a case enactment designed to be studied from the

basic models of psychopathology and treatment. This fine film may be ordered separately or in a series of films displaying various representative scholars discussing the hypothetical Otto from their own model of therapy.

The paper-and-pencil materials include three worksheets: (a) Intake and Problem Formulation, (b) DSM-III Classification, and (c) Treatment Plan and Prognosis. These forms should be reproduced for each student.

INSTRUCTIONS

Undergraduate students are assigned at random to one of several "clinic" groups representing a model or school of psychopathology (e.g., psychoanalytic, behavioral, client-centered, existential, biological). We have found that groups of 6-10 students function effectively for these exercises. During the course of the semester, as clinics are convened, students are rotated to other models. This procedure ensures that each student is exposed to the full complement of therapeutic approaches and helps each student develop ease of communication with other members of the class.

The clinic simulation procedure is generally introduced approximately 4 or 5 weeks into the semester. This timing permits the instructor to cover, in standard sequence, introductory chapters on abnormal psychology, the various models or approaches to psychopathology, the chapters on classification and assessment procedures, and at least one content chapter focusing on the disorders themselves. These chapters are usually followed by material covering the remaining diagnostic categories. The experienced instructor will frequently find that most textbooks follow this fairly standard sequence.

Critical to the success of the exercise is the effective introduction of the material in chapters about models of psychopathology and the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III)* (American Psychiatric Association, 1980). When these chapters are taught (and they should be introduced early in the course), it is important to prepare the students for the eventual clinic exercise by demonstrating how a case can be viewed by the various models, regardless of the diagnosis agreed upon from the *DSM-III*. For example, depression can be viewed as resulting from a chemical imbalance, from a reduction in reinforcement, or from a symbolic loss accentuated by a fixation in the oral stage of development. Thus, before students are actually given the clinic assignment, it is important that they have understood to some degree the ideas behind the various models, the nature of the classification system, and at least one area of disorders outlined in the *DSM-III*.

Time requirements for the exercise vary with the depth of each instructor's involvement with the procedure. For example, some instructors may not wish to have the clinic groups complete each form for each case presented. However, after the stimulus case has been presented, at least 30 minutes of discussion and processing should be made available for the groups. Additional time for a general review of each group's findings, a discussion of agreement within and among groups, and a presentation of new material by the instructor is needed. For thorough processing of the exercise, a full hour should be allotted. Group discussion and the processing of information can be carried over from one class session to the next. However, we have found the process to be smoother when the entire exercise is completed within one class session. Problems of absenteeism, retention of the case material, and so forth are minimized when closure is achieved within one class period.

Students in each of the miniclinics are presented with an overview of the nature of the learning task. Prior to the actual presentation of case material, students in each clinic are reminded of the particular concerns and variables relevant to their dis-

pline. The psychoanalytical group is, for example, cued to look for important features of early childhood, for current defense mechanisms employed, and for the possible forms that an eventual transference might take. Similar coaching is directed toward members of other miniclinics. These instructions are provided to students in a general session so students can become further aware of the differentiation in task and approach of the various therapeutic schools.

When clinic assignments and the general review have been completed, the full case presentation is made, starting with background information provided by the instructor. Each group also receives a copy of the written case history. It is helpful to select cases with a good deal of background information so that students from each clinic have enough potential information to make a case for the relevance or efficacy of their mode. After the background information is presented, the recording of the case is played. Next, students divide into clinic groups to discuss the case and to complete the paper-and-pencil assignments. While the students are working in their groups, it is helpful for the instructor to interact with each group in progress.

The paper-and-pencil assignments consist of three worksheets: Intake and Problem Formulation, DSM-III Classification, and Treatment Plan and Prognosis. Students are instructed to tailor their observations to their assigned perspective on the Intake and Problem Formulation Worksheet. For example, under the Historical Antecedent section, behaviorists are encouraged to note possible early learning histories, psychoanalysts are told to report possible developmental trauma, and students in the physiological clinic are directed to focus on possible early signs of neurological disorder, brain injury, congenital problem, and so forth. Each student is required to complete this form, although group discussion may take place before each student has completed the assignment.

The DSM-III worksheet requires students in each of the clinics to review all five axes of the *DSM-III*. Of course, students are informed that they will rarely utilize all axes in classifying a case. As with the Intake and Problem Formulation Worksheet, each student is required to complete a form after group discussion. Students are encouraged to stick to their guns even if others in the clinic disagree. Because of this practice, reliability estimates may be made for each clinic.

The classification exercise is followed by the completion of the Treatment Plan and Prognosis Worksheet. It is here that students are able to exercise the most creativity and to display familiarity with their assigned model of practice. Interventions should be justified on the basis of problem documentation and relevance to the student's particular mode. The prognosis section is also to be completed with reference to the assigned model and with consideration of the available resources and circumstances relevant to the model (e.g., How well would an older adult with an IQ of 75 do in traditional psychoanalysis?).

With the worksheets completed, the full class reassembles to process the results of the exercise. The instructor leads a discussion of each worksheet activity, highlighting the points of view for each model represented. In addition, a discussion of agreement or disagreement on diagnosis, etiology, and treatment both within and among groups may be held. Rough estimates of percent agreement within groups may be calculated by simply dividing the number of agreements (with the correct diagnostic category) by the combined number of agreements and disagreements within each group. It should be made clear to students, however, that this number is only a rough estimate that is probably inflated because it does not account for the number of agreements that would be expected by chance. It is probably not worth the

time it would take to labor through a complete explanation of the probabilities of chance agreement for each diagnostic category.

DISCUSSION

The miniclinic method presents several advantages to the instructor as well as a few potential stumbling blocks. Advantages include teaching students the rigors of keeping within theoretical models, teaching the logical connection between an etiological conception of problems and interventions associated with particular models, and exploring the problem of how models relate (or do not relate) to the current classification system. Finally, students are able to learn clinical material without rote. They learn that the *DSM-III* is a real tool and enjoy exercising their diagnostic skill through safe risk taking in the miniclinic.

The problems associated with this process include the always-present possibility that passive students will remain passive and let others complete the exercise. The medical students' disease syndrome and the phenomena of the "instant expert" can also be stimulated by this activity, yet early warnings and effective feedback in class can do much to prevent these problems.

One final and nontechnical note summarizing the process is in order: Both students and instructors who use this process find it a lot of fun!

REFERENCES

- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- Benjamin, L. T., & Lowman, K. D. (Eds.). (1981). *Activities handbook for the teaching of psychology*. Washington, DC: American Psychological Association.
- McKeachie, W. J. (1978). *Teaching tips: A guidebook for the beginning college teacher* (7th ed.). Lexington, MA: Health.
- Radford, J., & Rose, D. (Eds.). (1980). *The teaching of psychology: Method, content and context*. New York: Wiley.

SUGGESTED READING

- Spitzer, R. L., Skodoe, A. E., Gibbon, M., & Williams, J. B. W. (1981). *DSM III casebook: A learning companion to the diagnostic and statistical manual of mental disorders (third edition)*. Washington, DC: American Psychiatric Association.

Intake and Problem Formulation Worksheet

Name: _____

Date: _____

Section: _____

Case: _____

Clinic assigned: _____

Presenting problem:

Historical antecedents:

Current observations:

Notes, Ideas, and Questions for Discussion

DSM-III Classification Worksheet

Name: _____
Date: _____
Section: _____
Case: _____

Clinic Assigned: _____

AXIS I: Disorders Usually First Evident in Infancy, Childhood or Adolescence
Assessment data:

DSM-III
classification: _____

AXIS II: Personality Disorders
Assessment Data:

DSM-III
classification: _____

AXIS III: Physical Disorders and Conditions
Assessment Data:

DSM-III
classification: _____

AXIS IV: Severity of Psychosocial Stressors (1-7)
Assessment Data:

DSM-III
classification
(1-7): _____

AXIS V: Highest Level of Adaptive Functioning in Past Year (1-7)
Assessment Data:

DSM-III classification
(1-7): _____

Notes, Ideas, and Questions for Discussion

Treatment Plan and Prognosis Worksheet

Name: _____
Date: _____
Section: _____
Case: _____

Clinic Assigned: _____

Summary of needs (conditions):

Unanswered questions and further assessment needed:

Recommended interventions (with rationale):

Prognosis (related to specific problems):

Notes, Ideas, and Questions for Discussion

Discovering the Relationship Between Operational Definitions and Inter-observer Reliability

15 DISCOVERING THE RELATIONSHIP BETWEEN OPERATIONAL DEFINITIONS AND INTEROBSERVER RELIABILITY

Angela H. Becker

Indiana University Kokomo

This activity is designed for students in research methods and behavior modification classes or the methods section of other content courses. Students observe a brief videotape and collect data on the occurrence/nonoccurrence of a series of six behaviors. The main purposes of this activity are to help students (a) understand the importance of having clear operational definitions, (b) learn to calculate interobserver reliability, and (c) think about ways to improve a study that has low interobserver reliability. In addition, students gain practice in using time sampling and come to realize that observation as a data collection technique is more complex than casual observation.

CONCEPT This exercise helps students realize that observation as a data collection technique is more complex than casual observation. It introduces students to the use of time sampling, the calculation of interobserver reliability, and the importance of having clear operational definitions.

MATERIALS NEEDED You will need a 10-min videotape of human or animal behavior, a VCR and monitor for showing the tape to the class, a watch with a second hand, and enough copies of the handouts described later for each student in the class. Students will need pencils and will probably want calculators.

To be most effective, the videotape should be of a group of humans or animals that are active enough to produce several different types of behaviors. If at least some of those behaviors occur quite frequently and in several individuals at a time, students will come away with a better understanding of why time sampling is useful. My tape is of a group of white geese at a local park. (I would be happy to provide a copy of this taped segment to anyone who sends me a blank videotape.) There are many other possibilities for footage that will meet the previously described requirements. For example, you could videotape small children performing at a school program or playing at a birthday party, or you could get some footage of one of the livelier species at your local zoo. If you or someone you know is planning to visit another country, you may be able to obtain a tape of a festival or other group event from another culture. The videos that some universities make of their graduation ceremonies could also be used. If you have access to a VCR, you could tape a segment of an appropriate televised event. Although many television programs show groups engaging in behaviors that meet the criteria identified at the beginning of this discussion, most do not show this activity, uninterrupted by close-ups of individuals or pans to scenery and other locations.

in the story line, for more than a minute or two. There are exceptions, however, that would make good tapes for this activity: Televised New Year's Eve bashes usually show quite long segments of partyers, sporting events such as basketball or volleyball also show fairly long segments of activity on the court, and dance club shows on cable channels run segments of couples dancing for the duration of a complete song.

Prepare three handouts. The first should contain a data collection sheet with a row for each observation interval and a column for each behavior students are to record (see appendix A). The second should contain a list of behaviors (I recommend no more than six to eight), their operational definitions (some of which are purposely clearer than others), and a simple formula for inter-rater reliability (see appendix B). The third handout should contain a set of postobservation questions (see appendix C).

INSTRUCTIONS

This activity should be prefaced with a lecture on the use of observation techniques, including the advantages and disadvantages of time sampling in relation to other observation techniques (e.g., event sampling and narrative recording) and the concept of interobserver reliability. Although the basic observation techniques described in methods textbooks are much the same, the labels given to particular techniques vary. The following definitions of observation techniques are provided to facilitate gathering background lecture material and to prevent misunderstanding.

Time sampling is a technique in which the observer defines several target behaviors, divides the observation period into short intervals, and then alternates from observing to recording every other interval. In contrast, in *event sampling* the observer defines a target behavior and records every instance of that behavior as it occurs throughout the observation period. A *narrative recording* is a running description of behavior in which everything that is said or done during the observation period is recorded. The following are particularly useful sources for lecture material: chapter 6 from Bordens and Abbott (1996); chapters 6, 7, and 8 from Martin and Bateson (1993); and chapter 19 from Martin and Pear (1996).

Give each student all three handouts, and allow them a few minutes to read through the list of behaviors and operational definitions and become familiar with the layout of the data collection sheet. Have students work in 15-s intervals—alternating between 15 s for observing and 15 s for recording observations completed in the previous 15-s interval. For each observation interval, they simply look for whether or not each target behavior occurs. If the behavior occurs *at least once* in the observation interval, they are to place a tally mark in the appropriate column on the data sheet during the recording interval. Rather than having students keep track of their own intervals, use a watch with a second hand to time intervals for them. It works best if you simply call out "observe" or "record" at the beginning of alternate 15-s intervals. Explain to students that data collection will last for a total of 10 min. Each minute represents an observation period and is divided into 30-s sessions. During each 30-s session, students will have a 15-s interval to observe the behavior on the videotape and a 15-s interval to record their observations. It is important for students to understand what they are observing and recording; be sure to explain that they are recording the occurrence of target behaviors. That is, they are looking for whether or not a behavior occurs; they are *not* looking for the number of times a specific behavior occurs.

After 10 min, have students stop observing and work on the postobservation questions (see appendix C). As appendix C illustrates, students will first answer several questions individually, then compare those answers with a partner, and finally calculate interobserver reliability with their partner for each of the target behaviors.

DISCUSSION

Follow up with a class discussion of students' responses to the postobservation questions. Focus on those questions where partners' responses differed most often and on those behaviors that had the highest and lowest interobserver reliability. Then discuss possible reasons for these trends. With my tape of geese and set of behaviors, for example, students generally have very high interobserver reliability for displays and for tail shakes and low reliability for feeding and submission. When exploring possible reasons for these findings, student comments tend to focus on the importance of careful operational definitions and on problems with observation. For example, students notice that my operational definition of display behavior is much more concrete than my definition of submission, that the definition for feeding was too narrow to encompass much of what they wanted to be able to code as feeding behavior, and that several of the behaviors were difficult to identify accurately because of the distance from which the videotape was shot. For example, one student declared that she wanted a better definition of feeding, because "sometimes I thought they might be, but I couldn't see if they really had food in their mouths or not."

Next ask students to offer possible solutions for the reliability problems they have encountered. We talk about clarifying operational definitions. For example, several students decided that "touching beak to the ground several times in a row" would have defined feeding in a way that would have allowed them to record what they thought was feeding behavior. Students also brought up the possibility of reviewing the tape and discussing discrepancies between observers in order to resolve disagreements or practicing with sample tapes to improve reliability before viewing actual data tapes. This second idea was an elaboration on one student's comment that he "wished we could have watched the whole tape first while reading the definitions and *then* done the recording part." The students also decided that the value of a high-power zoom lens should not be underestimated if one wants to observe detailed behaviors and remain unobtrusive. Overall, students' responses to the activity indicated that not only did they learn a great deal, but that they enjoyed the activity as well.

A minor variation on this activity could allow students to discover for themselves that one of the pitfalls to time sampling is that there will always be lost data (i.e., behaviors that occur during recording intervals rather than observation intervals). Instead of having the entire class observe and record during the same intervals, divide the class in half, and have each half observe and record during opposite intervals. Have each half of the class pool their data and calculate the mean number of intervals in which tail shakes, feeding, grooming, display, aggression, and submission were observed by their group. Students should find that for behaviors that occur frequently, there will be little difference between the means reported by each half of the class. For example, the two halves of the class should be quite similar on mean number of tail shakes, simply because this occurs almost continuously among geese. However, for relatively infrequent behaviors, such as displays of aggression, students are likely to notice differences between

WRITING COMPONENT

reports by the two halves of the class. This can lead to a discussion of the relative usefulness of time sampling versus event sampling for observing infrequent behaviors. (Obviously, if you use this alternative procedure and you still want students to calculate interobserver reliability, they must do so by pairing up with someone from the same half of the class.)

Instructors who want to provide their students with an opportunity to do more writing than the small amount required to complete appendix C may add one of the following writing components to the activity.

1. Have students reflect on their expectations of observation in general and time sampling in particular. After giving students a brief description of the exercise they are about to engage in, ask them to respond to the question, "What do you think will happen when we do this time sampling observation?" After the exercise is complete, have students reread their earlier expectations and write a response to the following two questions: (a) "Which of your earlier expectations were met and why do you think this happened?" and (b) "Which of your earlier expectations were *not* met and why do you think this happened?" As a follow-up, students could construct a list on the chalkboard of the group's most common expectations, identify those that were not met, and then discuss whether those unmet expectations would make them more or less likely to want to use this method in their own future research.
2. Have students write a report to the researchers who set up the study. In that report, students should point out the strengths and weaknesses of the study and suggest improvements. This writing component could be followed by a small-group discussion in which students compare the strengths and weaknesses they noticed and try to identify the most methodologically sound and practical suggestions for improvement. You could also ask the students to use this small-group time to rewrite the operational definitions that they found lacking.
3. Instructors who have their students keep journals might consider having them include an entry about this observation activity. Students could be asked to respond to the question "What do you feel you learned from this observation exercise?" If content analysis is covered in your course, you could have students use these journal responses as data and attempt to code them into categories.

REFERENCES

- Bordens, K. S., & Abbott, B. B. (1996). *Research design and methods: A process approach* (3rd ed.). Mountain View, CA: Mayfield.
- Martin, G., & Pear, J. (1996). *Behavior modification: What it is and how to do it* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Martin, P., & Bateson, P. (1993). *Measuring behavior: An introductory guide* (2nd ed.). Cambridge: Cambridge University Press.

Appendix A

Data Collection Sheet for Time Sampling

Minute	Tail Shake	Feeding	Grooming	Display	Aggression	Submission
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

Each cell represents a 15-s observation interval. Recording intervals are not shown on this sheet.

Appendix B

Target Behaviors and Operational Definitions

Tail shake	Flicking tail back and forth rapidly several times in succession
Feeding	Actually taking food in beak
Grooming	Preening—using beak to fluff or pick at feathers
Display	Full (or almost full) extension of wings accompanied by several flaps, a slight lift in body posture, and a slight extension of neck—usually done when standing or when walking very slowly
Aggression	Nipping or threatening (by chasing or quickly swinging head toward another individual)
Submission	Running from or obviously avoiding close contact with another individual

Formula for Calculating Interobserver Reliability When Doing Time Sampling

$$\text{reliability} = \frac{\text{agreements}}{\text{agreements} + \text{disagreements}} \times 100$$

(Agreements = number of intervals in which you both marked that the behavior occurred, and disagreements = number of intervals in which only one of you marked that the behavior occurred.)

Appendix C

Postobservation Instructions

1. Individually, tally the number of intervals in which each behavior occurred. Then, answer the following questions:
 - Which behavior occurred *most* often?
 - Which behavior occurred *least* often?
 - Are there any behaviors that at least *appear* to be highly correlated? (That is, are there any behaviors that seem to always, or almost always, occur during the same intervals?)
2. Pair up with another student and compare your answers to the preceding questions. Did you disagree on any of them? If so, which one(s)? *Why* did you disagree?
3. Calculate the interobserver reliability between you and your partner for each behavior category. Identify the category that has the highest interobserver reliability and the category that has the lowest interobserver reliability.

Tail shake =

Display =

Feeding =

Aggression =

Grooming =

Submission =

Information Processing Capacity: A Visual Demonstration of the Magical Number Seven

INFORMATION PROCESSING CAPACITY: A VISUAL DEMONSTRATION OF THE MAGICAL NUMBER SEVEN

Fairfid M. Caudle

The College of Staten Island, City University of New York

This activity provides a visual demonstration of the well-known limitation on information processing capacity represented by the G. A. Miller's phrase "the magical number seven, plus or minus two." Students are presented with arrays of dots, arranged either randomly or in patterns. A graph of students' judgments of the number of dots in each array demonstrates the limits of information processing capacity and the facilitative effect of chunking. The demonstration also provides opportunities to explore aspects of experimental design and descriptive statistics.

CONCEPT The phrase "magical number seven, plus or minus two" refers to the limited capacity of short-term memory (Miller, 1956). This activity, unlike the auditory demonstrations typically included in introductory texts, uses visual stimuli to demonstrate this capacity and the value of chunking. It easily can be extended to cover experimental design and descriptive statistics.

**MATERIALS
NEEDED** In addition to chalk and a chalkboard, you will need 17 stimulus items, each constructed from a sheet of 8½-by-11 in. white paper and black or blue colored adhesive dots approximately $\frac{3}{4}$ in. in diameter. These dots are available in office supply stores. A total of 136 dots are needed.

Prepare the 17 stimulus items as indicated in appendix A. Each must consist of one sheet of paper with the number of dots indicated distributed either randomly or in a pattern. On the back of each sheet note lightly, for your own reference, the stimulus item number as well as the number and distribution of the dots. This will enable you to check that the sheets are in the proper sequence before beginning the demonstration.

The construction of stimulus items as described here has proved adequate for classes of up to 90 students. For larger classes, you may want to construct the stimulus items by placing adhesive dots directly on overhead transparency sheets. These, however, are more difficult to handle. If you use this method, check in advance to make sure all the dots on each transparency actually project on the screen.

You may also want to prepare a summary sheet that indicates the random stimulus item numbers in order of increasing number of dots (stimulus items 6, 3, 12, 1, 16, 14, 9, 11, 5, 4, 17, 2, 8, and 13) and the patterned item numbers in

the same order (stimulus items 15, 7, and 10). This will facilitate the construction of the results graph.

INSTRUCTIONS *Preparation*

On a table in front of you, arrange the stimulus items face down from Item 1 (on top) through Item 17 (on the bottom). Ask the class to turn to a blank page in their notebooks and number from line 1 through line 17.

Say to the class: "I am going to show you some sheets of paper with dots on them. For each sheet, I will give you three beats to get ready, one beat to watch, and one beat to write down the number of dots you see on that sheet. Write your answer for each sheet on a different line, going from Line 1 through Line 17. I will not be calling out line numbers, so just keep going until we finish. For each sheet, I will say, 'dah, dah, dah, look, write.'" Demonstrate with hand motions how you will hold up a stimulus item on "look."

Stimulus Presentation

Show each sheet by counting, at approximately 1 sec per beat, "dah, dah, dah, look, write." As you say "look," hold up a stimulus item. As you say "write," put the sheet face down. Repeat for each stimulus item.

RECORDING RESULTS

Draw a graph on the chalkboard. Label the vertical axis "Number of Persons Correct" and mark it in units of 10. Label the horizontal axis "Number of Dots" and number it from 1 to 14 (the maximum number of dots).

Referring to your summary sheet listing all the stimulus items, go through each item in the order of increasing dots and ask for a show of hands as to the number of students who wrote down the correct number. For example, say, "Sheet 6 had one dot. How many of you were correct?" Follow this with, "Sheet 12, two dots," and so on. It is helpful to know in advance how many are actually present so you can subtract the number of people wrong when almost everyone is correct. With a large class, divide the class into sections and have someone count each section.

Using this procedure, record the number of persons correctly responding to random arrangements of 1 through 14 dots. For each stimulus item, count the number of correct persons and plot a solid dot at the appropriate place on the graph. Connect the solid dots with solid lines to complete the graph. Then record the number of persons correctly responding to the three items with dots arranged in patterns. Indicate the number correct with small hollow circles and connect these with broken lines. Complete your graph with a key indicating that solid circles connected by a solid line correspond to random arrangements, whereas hollow circles connected with a broken line indicate pattern arrangements.

DISCUSSION

Typically, for random arrangements of dots, virtually the entire class is correct for 1 through 5 dots. Thereafter, the number of persons correct begins to decline, and does so precipitously for 10 through 14 dots. Your graph will not be perfect, but the trend should be clearly apparent.

Once you have constructed the graph for random arrangements, ask the class to interpret the graph. Identify the point where lots of people begin to make

WF
Co

mistakes and relate this to the "magical number seven," which represents our information-processing capacity.

When dots are arranged in patterns, the number correct is always higher. Ask the class to compare the number correct for 9, 10, and 12 dots arranged in patterns with the number correct for 9, 10, and 12 dots distributed randomly. Ask for suggestions as to why the results are as they are. Introduce the concept of chunking to explain the dramatic increase in capacity when information is organized into patterns.

After discussing the main findings of the demonstration, you can extend the activity by having students analyze it in terms of experimental design. Have them identify the independent variables (there are two: the total number of dots and the type of arrangement, random or pattern) and the dependent variable (number of persons correct) for the demonstration. Ask for someone to state a relationship between the independent and dependent variables that was illustrated by this activity. (As the number of dots increased, the number of persons correct decreased. However, the number of persons correct was higher when the dots were in a pattern.)

Continuing your discussion of variables, ask the class to identify possible uncontrolled variables that might have affected the outcome. These might include such things as distance from the stimulus items, viewing angle, movement of stimulus items as you held them up, inadvertent variations in viewing time, and so forth. Ask for suggestions as to how to adequately control for these variables, and describe laboratory instruments, such as the tachistoscope, that have been designed to enable increased control over such variables.

Finally, it is good to point out how a very large number of individual responses (roughly 17 times the number of people who participated) can be summarized by means of a single graph. This illustrates the value of graphs and other forms of statistics for making data manageable and understandable.

WRITING COMPONENT

Several writing exercises can be assigned to assess students' understanding of information-processing capacity as illustrated by this activity. For example, ask the students to write a paragraph summarizing the purpose of the demonstration and describing the independent variable, the dependent variable, and the relationship between them.

As a follow-up to the classroom discussion of uncontrolled variables, you can also have students choose one of these variables and design an experiment in which it becomes an independent variable while other variables are controlled. Have students explain in writing how results could be summarized in a graph.

Finally, have students write descriptions of situations in which chunking of visual information into patterns is important. These might include occupations (e.g., air traffic controller or musician), sports (e.g., football or basketball), board games (e.g., chess), and activities of daily life (e.g., finding one's car in a large parking lot or finding items during a trip to a supermarket). In each instance, have students indicate why they believe chunking is of value. Some examples of studies reporting visual chunking are noted in the Reference and Suggested Reading sections.

REFERENCE

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

MEMORY

**SUGGESTED
READING**

- Allard, F., & Burnett, N. (1985). Skill in sport. *Canadian Journal of Psychology*, 39, 294-312.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Cohen, G. (1989). *Memory in the real world*. Hillsdale, NJ: Erlbaum.
- Squire, L. R. (1992). *Encyclopedia of learning and memory*. New York: Macmillan.

Appendix A

Stimulus Items

Item Number	Number of Dots	Distribution
1	4	Random
2	12	Random
3	2	Random
4	10	Random
5	9	Random
6	1	Random
7	10	Random
8	13	Pattern (2 rows of 5)
9	7	Random
10	12	Random
11	8	Pattern (4 rows of 3)
12	3	Random
13	14	Random
14	6	Random
15	9	Random
16	5	Pattern (3 rows of 3)
17	11	Random

The Role of Prior Information in Dream Analysis

80 THE ROLE OF PRIOR INFORMATION IN DREAM ANALYSIS

Douglas A. Bernstein

University of Illinois

In this activity, the class receives a handout describing the demographic characteristics and brief life history of a woman named Doris, along with a dream she reported having. At the bottom of the handout, there is space for the student to interpret the meaning of Doris's dream. The dream is the same on each handout, but the description of Doris is not.

CONCEPT This activity is designed to help students understand that the interpretation of dreams, like the interpretation of other stimuli, can be influenced by prior knowledge, expectancies, motivation, emotion, and other top down processes. More specifically, it illustrates how easily a clinician's prior knowledge about a client might prompt very different conclusions about the meaning of dream content.

This activity provides an easy way to show students the link between principles that guide the perception of objects and those that operate in social perception in general and in psychotherapy in particular.

MATERIALS NEEDED You will need copies of each of the three handouts given in appendixes A, B, and C. Collate the three items before you hand them out, so that one third of the students will receive each version.

To aid class discussion of varying dream interpretations, it is helpful to have each version of Doris's demographics and life history on overhead transparencies in print large enough to be read from the back of the room.

INSTRUCTIONS Give each student a version of the Doris handout. After a few minutes, ask the students to write their interpretation of the dream at the bottom—and perhaps the back—of the handout. You can either collect and read aloud some of the interpretations, or just ask students to read them aloud.

DISCUSSION As the readings proceed, it will soon become obvious to the class that something is amiss. At this point, reveal the differences in the three descriptions of Doris and point out the influence those differences had on the interpretations of Doris's dream.

You can end this demonstration by pointing out that clinicians are aware of the role of prior knowledge in dream analysis (and other aspects of therapy). This is why, for example, they tend to base conclusions about clients on a series of dreams rather than on just one and why they seek to combine assessment information from various sources, such as tests and interviews. I have found that stimulating class discussions result from pointing out that, in spite of such efforts, clinicians are as vulnerable as the rest of us to the biasing effects of expectancy in dealing with clients.

**WRITING
COMPONENT**

In addition to having students write their interpretation of the dream in class, you might also ask students to write a summary of the purpose of the demonstration and what they learned from it. This can be done in 5 min at the end of the exercise or at the end of class. A quick perusal of the summaries can show if students understood the point that you were trying to make. Summaries are efficient ways to check for student understanding.

REFERENCE

Ullman, M. (1986). Access to dreams. In B. B. Wolman & M. Ullman (Eds.), *Handbook of states of consciousness*. New York: Van Nostrand Reinhold.

C Further Reading

DSM-III

- Bayer, R. and R. Spitzer (1982). "Edited Correspondence on the status of homosexuality in DSM-III." *Journal of the History of Behavioral Sciences* 18(1): 32–52.
- Bieber, I. (1962). *Homosexuality: A Psychoanalytic Study*. New York, Basic Books.
- Drescher, J. and J. P. Merlino (2007). *American psychiatry and homosexuality : an oral history*. New York, Harrington Park Press.
- Evans, A. (2010) "Zap, You're Alive!", <http://gaytoday.badpuppy.com/garchive/viewpoint/083099vi.htm>
- Fryer, J. (2008). "Dr. Anonymous Speech." Retrieved 10/6/2008, 2008, from http://www.oism.info/en/society/social_control/dr_anonymous_apa_speech_1972.htm.
- Gold, Ron (2004). 'Three Stories of Gay Liberation' <http://www.queerstories.org/custom.html>.
- Moor, P. (2002). "The View from Irving Bieber's Couch: 'Heads I Win, Tails You Lose'." *Journal of Gay & Lesbian Psychotherapy* 5(3/4): 25–36.
- Nicolosi, J. (1991). *Reparative Therapy of Male Homosexuality*. Jason Aronson, Inc., , Northvale, New Jersey.
- Nicolosi, J. (1993). *Healing Homosexuality: Case Stories of Reparative Therapy* Northvale, NJ, Jason Aronson.
- Scasta, D. L. (2002). "John E Fryer, MD, and the Dr. H. Anonymous Episode." *Journal of Gay & Lesbian Psychotherapy* 6(4): 73–84.
- Sheehy, N., A. J. Chapman, et al. (1997). *Biographical dictionary of psychology*. London ; New York, Routledge Reference.
- Socarides, C. W. (1963). "Review of Homosexuality. A Psychoanalytic Study : By Irving Bieber, et al. New York: Basic Books, Inc. 1962. 358 pp." *Psychoanalytic Quarterly* 32: 111–114.
- Society of Medical, P. and I. Bieber (1962). *Homosexuality; a psychoanalytic study*. New York, Basic Books.
- Silverstein, Charles The Origin of the Gay Psychotherapy Movement (Published in Duberman, A Queer World, New York University Press, 1997. 7,300 words <http://drcsilverstein.com/origin.html>.
- Spitzer, R. (1981). "The diagnostic status of homosexuality in DSM-III: a reformulation of the issues." *American Journal of Psychiatry* 138: 210–215.
- Spitzer, R. L. and American Psychiatric Association. (1981). *DSM-III case book : a learning companion to the Diagnostic and statistical manual of mental disorders (third edition)*. Washington, D.C., American Psychiatric Association.

Spitzer, R. L. and J. Endicott (1978). Medical and Mental Disorder: Proposed Definition and Criteria. Critical Issues in Psychiatric Diagnosis. R. L. Spitzer and D. F. Klein. New York, Raven Press: 15–40.

Stoller, R.J., Marmor, J., Bieber, I., Gold, R., Socarides, C.W., Green, R. and Spitzer, R. "A Symposium: Should Homosexuality Be in the APA Nomenclature?" *American Journal of Psychiatry* 130(11) p. 1207–16

Zusne, L. (1984). Biographical dictionary of psychology, Greenwood Press.

D Obituaries and Tributes

Anne Anastasi

"Anne Anastasi: Biography of a Former APA President" American Psychological Association: <http://www.apa.org/about/archives/presidents/bio-anne-anastasi.asp>.

Erica Goode. 2001. "Anne Anastasi, the 'Test Guru' of Psychology, Is Dead at 92". *The New York Times* (): <http://www.nytimes.com/2001/05/16/nyregion/anne-anastasi-the-test-guru-of-psychology-is-dead-at-92.html>.

George Albee

Kessler, M. (2007). "Obituary: George W. Albee (1921–2006)." *American Psychologist* 62(4): 317–318.

Associated Press (2006) "George W. Albee; Psychologist Sought Social Cures" July 13, 2006

Marie Skodak Crissey and George W Albee. 1982. "The 1982 division 27 award for distinguished contributions to community psychology and community mental health: George W. Albee". *American Journal of Community Psychology* 10, no. 1 (): 1–36

Richard H Price. 2007. "George Albee: A man with a better idea". *The Journal of Primary Prevention* 28, no. 1 (): 55–57

Margaret Gatz. 2007. "In memory of George Albee". *The Journal of Primary Prevention* 28, no. 1 (): 19–20

Irving Bieber

Myers, S. L. "Irving Bieber, 80, a Psychoanalyst Who Studied Homosexuality, Dies" NY Times 8/28/1991 <http://query.nytimes.com/gst/fullpage.html?res=9DoCE3D7153FF93BA1575BCoA967958260>.

Kenneth Clark

Phillips, L. (2002). "Recontextualizing Kenneth B. Clark: An Afrocentric Perspective on the Paradoxical Legacy of a Model Psychologist-

Activist." Evolving Perspectives on the History of Psychology. W. E. Pickren and D. A. Dewsbury. Washington, DC, American Psychiatric Association.

John Fryer

Pray, R. (2003). "John E. Fryer, 65, psychiatrist." The Philadelphia Inquirer. Philadelphia: Bo9.

Lenzer, J. (2003). "John Fryer" BMJ March 22; 326(7390): 662.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1125557/>.

Barbara Gittings

Kameny, Frank (2007). "Frank Kameny's Eulogy for Barbara Gittings" available at <http://www.kamenypapers.org/eulogy.htm>.

Harry Harlow

Sears, R. R. (1982). "Obituary: Harry Frederick Harlow (1905–1981)." American Psychologist 37(11): 1280–1281.

Obituary for Harry Frederick Harlow (1905–1981). Harlow's research career, spanning a full half century, was notable for its single-minded adherence to the search for understanding of a single species, for the imaginative methods he devised for the study of both cognition and motivation, and for the crucially important findings these methods provided. His careful empirical development of the facts of behavioral development led to prepositional generalizations that have contributed notably to an understanding of human as well as monkey motivational development. (PsycINFO Database Record (c) 2008 APA, all rights reserve

Harold Lief

Pearce, J. (2007). "Harold I. Lief, Advocate of Sex Education, Dies at 89", The New York Times, March 23, 2007: <http://www.nytimes.com/2007/03/23/obituaries/23lief.html>.

Charles Socarides

Fox, M. "Charles W Socarides, Psychiatrist and Psychoanalyst, Is Dead at 83", NY Times, 12/25/2005 <http://www.nytimes.com/2005/12/28/nyregion/28socarides.html>.

John P. Spiegel

Cook, J. "John P. Spiegel, 80, Expert on Violence And Combat Stress" New York Times, July 19, 1991. <http://www.nytimes.com/1991/07/>

19/obituaries/john-p-spiegel-80-expert-on-violence-and-combat-stress.html.

Kurt Freund

<https://www.nytimes.com/1996/10/27/us/kurt-freund-dies-at-82-studied-deviant-sexual-arousal.html?mtrref=www.google.com>
(paywalled)

Helmore, Edward 1996

Rob Gold

Andy Humm. 2017. *Ron Gold, Pioneer in Challenging Sickness Label, Dies.* <https://www.gaycitynews.nyc/stories/2017/10/w27290-ron-gold-pioneer-challenging-sickness-label-dies-2017-05-16.html>

