2018

Training Games: An Application of Game Theory to Clinical Psychology Graduate Training

Benjamin David Cornell
University of Denver

Training Games: An Application of Game Theory to Clinical Psychology Graduate Training

Abstract

Contents

Document Type

Capstone

Degree Name

Psy.D.

Department

Graduate School of Professional Psychology

First Advisor

Michael Karson

Second Advisor

Peter Buirski

Third Advisor

Mac Kentnor

Keywords

Game theory, Clinical psychology, Graduate students

Training Games:

An Application of Game Theory to Clinical Psychology Graduate Training

A DOCTORAL PAPER
PRESENTED TO THE FACULTY OF THE
GRADUATE SCHOOL OF PROFESSIONAL PSYCHOLOGY
OFFICE OF GRADUATE STUDIES
UNIVERSITY OF DENVER

IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE
DOCTOR OF PSYCHOLOGY

BY
BEN CORNELL
April 16, 2018

APPROVED: _____
Michael Karson, Ph.D., J.D., Chair

_____
Peter Buirski, Ph.D.

_____
Mac Kentnor, Psy.D.

to Anna and Michael, for obvious reasons

"All I do is separate the game from the truth."

— *The Notorious B.I.G.*

Contents

Groundwork[1]

Game Theory

Game theory is a conceptual framework for studying the behavior of participants in "situations of conflict and cooperation between rational decision makers" (Tadelis, 2013, p.xi) as well as those situations (called "games") themselves. Specifically, game theory is concerned with "strategic" decision-making, meaning decisions made by one participant which take into account the behavior of other participants in the system. To illustrate:

> "Think of the difference between the decisions of a lumberjack and those of a general. When the lumberjack decides how to chop wood, he does not expect the wood to fight back; his environment is neutral. But when the general tries to cut down the enemy's army, he must anticipate and overcome resistance to his plans" (Dixit & Nalebuff, 1993, p. 1).

As such, Game Theory's games are by definition "multi-agent" systems, in that they involve more than one "decision maker" in interaction.

To a game theorist, a game is often laid out in a decision tree, (or more commonly a grid, but a tree is a better visual) where a branching set of successive decisions and corresponding responsive decisions lead from each player at the start of a game to a set of possible "outcomes" or "payouts" at the outermost branches of the tree. Different games have different characteristics, some of the most important of which are whether they are zero-sum or non-zero-sum; whether moves are made sequentially or simultaneously; whether players possess complete

[1] **A Note on Terminology**
Calling such interpersonally intimate activities as psychotherapy and supervision of psychotherapy "games" has a decidedly cynical ring to it, especially at first glance. Within the mathematical and logical discipline called "Game Theory," a "game" is a situation with a specific set of characteristics, but referring to a situation as a "game" says nothing about its emotional significance. Similarly, calling an individual's decision-making "strategic" only implies that it takes other agents into account, but not (necessarily) that it is especially selfish, cold, or manipulative.

or incomplete information about one another; and whether or not players can communicate with one another or learn from their opponent's past behavior (Straffin, 2004). The characteristics or rulesets of different games lead players to play them differently, with different sorts of games preferencing different sorts of strategies.

As an academic discipline, Game Theory was originally focused on formal abstractions of conflict and cooperation, with well-defined rules, choices, and outcomes. Over time, though, its applicability has proven to be surprisingly broad, with substantive contributions to such varied fields as economics, social psychology, and evolutionary biology. In abstract games, things like behavioral probabilities, and the precise point at which strategies become "stable" or optimal, are represented with variables, such that they can be calculated with perfect accuracy given numerically precise input parameters (Tadelis, 2013). When game theory is used as a lens through which to understand phenomena in the natural or social world, it is primarily an application of principles derived from the play of abstract games rather than a calculation of precise quantities or formulae. In order to apply those principles, the game theorist articulates bounded "games" in the environment, the "rules" of which are the practical constraints naturally encountered by a player in the game under consideration. Just as a first date could equally be seen as an "emotional transaction" or a "struggle for dominance," so too could it be understood as a "game."

The players in game theory's games are assumed to be "rational," meaning that each player's behavior serves to maximize her individual "payout." This is not to say that all the players *want* the same thing, but instead that all players behave in a way that (as far as the player knows) is most likely to result in what they, individually, want. The application of game theory to evolutionary biology is a particularly salient model in that evolutionary game theorists often

"work backwards" by starting with the observed behavior of organisms in the natural world and using these observations to understand the nature of the environment to which those behaviors were the optimal response (Dawkins, 1976). In other words, one can learn a lot about the context in which an organism evolved or developed by watching what it does, asking "what is the game for which these behaviors are the best strategy?" and then "finding" that game in the organism's environment.

With this working assumption in mind—that observed behavior is the individually optimal response to the strategic situation as perceived by the individual—we can infer the nature of the game a player is playing by observing what she does, just as we can infer what a rational agent will do if we completely understand the game in which she finds herself. So, considering psychology and game theory together, both the question "What game is this person playing optimally?" and "How would a person play this game optimally?" are relevant. When we encounter apparently irrational behavior, we can just as easily assume that we have missing information about the game perceived by the player, as to assume his behavior no longer conforms to "rationality." For example, even individuals with far higher hourly incomes than landscapers charge have been known to mow their own lawns, and this task is easily recognized as potentially rational because of the intangible mental "payoffs" it might provide such as a break from the kids, the satisfaction of physical activity, or reinforcement of an image of oneself as "the kind of person who does their own yard work."

Game Theory and the System of Clinical Psychology Graduate Training

Clinical Psychology, as its own system within the larger system of human society, presumes to provide the service of somehow ameliorating problematic behavioral and psychological functioning. Within this context, the discipline of psychotherapy is perhaps most

directly concerned with the task of affecting change in individual psychology. In the "Recognition of Psychotherapy Effectiveness," the American Psychological Association (2012, p. 102) defines "psychotherapy" as "the informed and intentional application of clinical methods and interpersonal stances derived from established psychological principles for the purpose of assisting people to modify their behaviors, cognitions, emotions, and/or other personal characteristics in directions that the participants deem desirable." The APA concluded that from the body of research comparing the effectiveness of various therapeutic modalities, "(1) most valid and structured psychotherapies are roughly equivalent in effectiveness and (2) patient and therapist characteristics, which are not usually captured by a patient's diagnosis or by the therapist's use of a specific psychotherapy, affect the results" (p. 103).

The effect of "therapist characteristics" on "the results" can potentially be quite large, such as that noted in Miller, Hubble, and Duncan's (2007) analysis of Wampold and Brown's (2005) study of client outcome across a sample of 581 mental health clinicians. They noted that "clients of the best therapists in the sample improved at a rate at least 50 percent higher and dropped out at a rate at least 50 percent lower than those assigned to the worst clinicians in the sample," and that "drugs used in combination with talk therapy were 10 times more effective with the best therapists than with the worst " (Miller, Hubble, & Duncan, 2007, p. 28).

While drawing substantive conclusions about the causal factors behind differences in client outcomes on a wide scale is generally difficult, clinician expertise seems to be a central contributor and the only one the therapist has more or less direct control over. The institution clinical psychology graduate training is precisely concerned with instilling this expertise; generally trying to teach aspiring therapists to do and say things that are more effective in encouraging positive psychological change than the alternatives. At least according to the APA,

client outcomes are at least partially determined by the therapists abilities, skills, and attitudes, and a therapist's abilities, skills, and attitudes are at least partially determined by her training.

So, at its most basic, the system of graduate training in clinical psychology concerns itself with conveying knowledge, skills, and credentials (for a fee) to students who are willing to pay for an opportunity to acquire them. Clinical psychology graduate training generally represents a somewhat cooperative task shared by training institutions and the students they train, with the ostensibly shared goal of producing competent practitioners with skills that will be valued in the economic marketplace and the "marketplace of ideas" of clinical psychology academia. Larkin (2015), for example, summarizes the goal of the training program as "creat[ing] well-socialized, ethical, and professional psychologists," (p. 304) though other training directors and program designers would doubtless describe what they are trying to produce in different terms. This overarching goal, and the "point" of the highest-level game in the system under consideration, orients all the other games being played in the system. This is not to say that in practice, the system might not be "pointed" at something else like financial gain, self-perpetuation, the psychological well-being of the larger society, or the advancement of particular political interests, just as the point of a chess game might be to develop the relationship between the players, entertain others, or sublimate aggression. But, some agenda like "creat[ing] well socialized, ethical, and professional psychologists" is the presumable answer to the question "What is this system *supposed* to be for?"

Psychological Payouts, Reconciling Game Theory and Clinical Psychology

In one sense, game theory and clinical psychology are exact opposites; the former concerns optimal play within a given ruleset and the latter concerns the infinite variation, development, and meaning of suboptimal behavior in humans. Comparing a relatively formal

game like chess to something like an admissions interview for a doctoral program in clinical psychology, it is immediately clear that the ruleset for the interview is far larger, reflective of the relative freedom of decision-making in an interpersonal playing field. But, both a chess player and an admissions-interviewee face a series of decisions between mutually exclusive alternative actions. The chess player cannot choose to both move her bishop and not move it on a given turn, and the interviewee cannot both tell a potentially-risky personal anecdote and not tell it. Both the chess player and the interviewee can be understood to be deciding between these alternatives primarily in reference to their desired outcome (something like "achieve a checkmate," on the one hand, and "convince this interviewer that I belong here" on the other).

At first glance, people routinely play the "games" they encounter suboptimally. For example, human beings make reliably skewed decisions when their consideration of alternatives involves very large, or very small statistical probabilities, and respond disproportionately to emotionally evocative eventualities (Kahneman, 2011; Wilson, 2002). In a very simple view, these phenomena could be seen as proof positive that human beings are *not* game theory's "rational agents," because they reliably behave contrary to the aim of maximizing their own payout, as all game theory's "players" do. The aforementioned chess player "misreads" the board and falls into traps, while the interviewee aims at displaying emotional vulnerability, but convinces his interviewer only that he has "poor boundaries." How do we make sense of the difference between "optimal" and "actually observed" here? Some of it is doubtless due to imperfect or ambiguous information and the limited processing power of the human mind, but the rest could be seen as the presence of something like "psychological payout modifiers" which modify the felt value of the available material or relational payouts. Operationally, we might

define the psychological component of outcomes as *the idiosyncratic meaning which the individual player assigns to each individual outcome on the payout table.*

Pattern recognition is one of the core functions of the human mind, at levels as basic as making sense of visual information, and picking out the boundaries of objects in the visual field, and as heady as the process of distilling emotional experience into "organizing principles" (Orange, Atwood, & Stolorow, 1997) that guide how all information is perceived and shape behavior outside of awareness. Fonagy *et al*, (2003, p. 416) describe a learning system that occurs in the first years of life, which result in,

> The creation of a processing system for the self (and significant others) in terms of a set of stable and generalized intentional attributes, such as desires, emotions, intentions, and beliefs inferred from recurring invariant patterns in the history of previous interactions. The child comes to be able to use this representational system to predict the other's or the self's behavior conjunction with local, more transient intentional states inferred from a given situation.

Importantly, the creation of this "processing system," or "system of organization," occurs in the particular child's individual developmental context, which will necessarily bear only limited resemblance to interpersonal contexts encountered later in the individual's life. Also, the individual will likely have only limited awareness that her predictions are based on that *particular* data set, and that those predictions may be therefore skewed when applied to novel environments. Brandchaft, Doctors, and Sorter (2010) note that:

> In the absence of sufficient self-awareness, the individual is blind to his role in structuring his own reality. The world in which he lives is experienced as though it were

> something independently and objectively real, rather than as something partly constituted by his own organizing principles (p. 52).

The individual understands situations based on his predictions about what will happen in them, what behaviors are likely to result in the most favorable outcomes, and how particular outcomes are likely to "feel." And, the particular way that specific outcomes feel is similarly subject to modulation through this same organizational structure.

It is useful to notice that—especially in looking at games played by humans in relational situations—a person's individual set of possible "payouts" in a given situation roughly constitute her "agenda." Whether she is making choices at a conscious level, cognitively weighing the risks and rewards of various actions, or responding unconsciously and toward ends which are not consciously symbolized, she "pursues" the *best-possible-outcome-all-things-considered,* as determined by her personal psychological organization. For example, agendas "pursued" by a student in a given hour of a psychological diagnosis course might include "sounding smart," "doing as little work as possible while maintaining an 'A'," and "tamping down an insidious doubt about being characterologically unsuited for psychotherapy and graduate school."

In those terms, game theory's assumption of "rational agents" is directly equivalent to Skinner's (1965) supposition that an organism's behavior is determined by the interplay of its reinforcement history and its environment. Relatedly, many of the theoretical sub-disciplines of historical and contemporary psychoanalysis tend to deal primarily with psychological behavior resulting from prior experience and "meaning making" (Buirski & Haglund, 2001), and interacting as a pre-constructed ruleset for subsequent encounters with "objects" (Mitchell, 1986) in the internal and external world.

The therapeutic encounter, conceptualized by Stolorow, Atwood, & Brandchaft (2014) "as an intersubjective process involving a dialogue between two personal universes" (p. 46) might yield a game theoretic understanding (i.e., the "game" inferred from the "behavior" observed in the interplay between analysand and analyst in a psychoanalysis) of something like "two agents, one of which pursues agendas such as 'feeling better,' 'impressing the analyst,' while the other pursues aims like 'inducing the analysand to speak openly about his thoughts,' and "getting the analysand to pay me." In this view, if one put real people with their own lifetime of memories into the formal games of game theory, one would expect their behavior to deviate from mathematically optimal play to the precise extent dictated by their idiosyncratic biopsychosocial makeup. Inversely, if we added our "psychological payout modifiers" to the rational agents in game theory, we would expect them to play like real people.

While game theory has an established tradition of application to social psychology, there is some disagreement about its compatibility with clinical psychology's understanding of human behavior. Flabbi and Pediconi (2014, p. 354) argue that this construct is incompatible with the classical psychoanalytic view of the drive-governed human unconscious:

> The law [of strategic interaction] established by game theory not only does not share the main features of the Freudian unconscious, but it is actually in fundamental opposition to it… Game theory describes interactions as implemented in a mathematical function: a given input maps always and immutably to a given output… It is the opposite of the concept of drive that, by definition, needs the other to be established in the subject."

Their contention is that the rational decision maker makes her decisions consciously in pursuit of her best payout, while the unconscious directs behavior toward the satisfaction of drives whose "aims" may well confound payout maximization.

However, in order for a "given input" to map to a known output, the characteristics of the input must be known completely. When looking at human behavior like a psychologist might, the "input" is generally something like *the comprehensive reinforcement history of the subject as well as that of all other agents in the system, as well as the specific "ruleset" of the game in question.* In that sense, Flabbi & Pediconi's drives that "need the other to be established in the subject" are implicit in the consideration of each individual's "payout scheme." One outcome will be more desirable than another due to the influence of the object of a drive, because "The principle of rational action still guides impulsive acts but as a function of available evidence about the 'pragmatic' aspects of a goal object, about the specific situational constraints on action, and about the dispositional constraints characteristic of the actor" (Fonagy et. al., 2003, p. 445).

A similar contention is that the strategic analysis of a situation neglects the situation's *meaning:* "While there is internal conflict in game theory. . . it is a conflict over making the right quantitative choice. It is a conflict of costs and benefits, of summing up quantitative amounts. What is principally missing from the intricately worked out calculus of game theory is meaning—especially, the meaning of the moves" (Alper, 1993 p. 52). It is exactly "the meaning of the moves" that the present discussion hopes to explore as we look at strategic behavior such as that employed by therapists and clients in psychotherapy or by individuals unconsciously selecting between a set of available "defense mechanisms" (Freud, 1937). The "choice" of meaning we assign to particular behaviors has substantial strategic significance in how it shapes the perceived game, the range of behaviors the individual "allows" herself and others, and the psychological and behavioral responses she makes to the strategic behavior of others.

If—in confronting the particular game of the doctoral paper requirement at the University of Denver's Graduate School of Professional Psychology's Psy.D. program—I respond by

writing about psychology and game theory and investing more time than would have been optimal to get the "rubber stamp" toward my degree, there are conclusions to be drawn about my conscious and unconscious agendas in the doctoral paper game, and about the nature of the game which I idiosyncratically perceive.

Perhaps, in addition to being a graduation requirement, I also see a playspace with opportunities to induce cherished mentors to say approving things about me. My choice to try to demonstrate the usefulness of viewing nuanced interpersonal interactions strategically may be an attempt to paint my pre-existing tendency to see things this way as "legitimate and valuable," rather than "vaguely distasteful." Choosing a topic that I suspected had a "decidedly cynical ring to it," may have been aimed at a payout of reinforcing a fantasy of myself as someone with "interesting and dissident ideas," and persuading others of the same. These choices and their corresponding hoped-for outcomes—whether consciously symbolized or not—carry the unique meaning I have made of myself in relation to the game at hand.

Alper's dismissal of meaning in strategic consideration mirrors the attitude that behaviorists and evolutionary biologists face when existing within the dominant framework of what Harari (2015) calls the "Humanist Religion." In Harari's language, "Religion is…any all-encompassing story that confers superhuman legitimacy on human laws, norms and values…by arguing that they reflect superhuman laws" (p. 211). The behaviorist and the biologist see humans as sites of often-conditioned behavior generation, while humanist dogma demands noncritical acceptance of the foundational truths that "*Homo sapiens* has a unique and sacred nature…that the unique nature of *Homo sapiens* is the most important thing in the world, and it determines the meaning of everything that happens in the universe" (Harari, 2011, p. 256). Because the behaviorist exists within a humanist hegemony, his means of inquiry (which

blaspheme the sacredness of the free-willed agent and meaning maker by offering "explanations" for behavior that arise from the environment rather than from self-determination) are labeled "cynical." To some degree, the psychotherapist may find herself in a similar position when suggesting that a client's behavior might arise from influences outside his awareness.

Games in the Training of Psychotherapists

Prisoners' Dilemmas, Internship and Graduate Applicant Selection

The bookends of a clinical psychology doctoral student's graduate training are the application process for graduate school at the beginning and the Association of Psychology Postdoctoral and Internship Centers (APPIC) national internship match at the end. These two contexts incorporate some of the most obviously competitive games in the graduate training process, as applicants and prospective interns compete for a finite number of desirable "places" at attractive programs, while graduate programs and internship sites compete for a finite pool of attractive applicants.

These games have many aspects in common with one of Game Theory's most famous games, "The Prisoner's Dilemma" (Dixit & Nalebuff, 1991). The classic formulation of this game is a situation in which two prisoners are facing sentencing and deciding whether to "snitch" on one another. Though the precise numbers in question aren't important, if the prisoners "cooperate" (with one another, not the hypothetical police) they get a relatively "good" outcome (say, one year in prison). If both prisoners "defect" (snitch on one another) they both get a "bad" outcome (say five years in jail, each). Finally, if one prisoner stays quiet and one talks, the one who defects gets a "best" outcome (immediate release) and the one who stays quiet gets a "worst" outcome (ten years). The "dilemma" part is that both know they will fare better

overall by cooperating, but each knows that in each situation they will do better individually by choosing to "defect." So even though they both get five years by defecting instead of one year for cooperating with each other, neither can pass up the strategic opportunity of immediate release, and thus doom themselves to the "second worst" outcome.

One of the interesting characteristics of this particular game is that it has what is known as a "dominant strategy." For a player in the prisoner's dilemma, the "defect" strategy is "better for him than all of his other other available strategies no matter what strategy or strategy combination the other player or players choose" (Dixit & Nalebuff, 2008 p. 70). In practical application to situations identifiable as "prisoners' dilemmas," the presence of this dominant strategy of defection means that players can be expected to be primarily motivated by self-interest rather than collective interest, except where idiosyncratic psychological organization incentivizes self-interest to coincide with collective interest. While the applicant who happens to prize honesty over career ambition might be inclined to disclose probably-invalidating information in an admissions interview, the "default" agenda of applicants in interviews is to present themselves in that light they perceive to be most compatible with the future outcome of an "offer letter."

The dominant strategy of defection for each applicant and each program considering *just this one application,* means it is nearly always strategically "wise" to attempt some degree of deception. If this game were played totally cooperatively by all parties involved, it might look something like this: A number of applicants apply to a number of graduate programs. In the admissions interviews, each applicant meets with a representative of each graduate institution, and during the meetings, the applicant and the representative describe to one another, with absolute honesty and perfect articulation, the qualities of the applicant and the graduate program

respectively (assuming they have perfect information about their own characteristics). Each pair arrives at a perfectly accurate estimate of the "fit" between the applicant and the program, and makes their admissions decisions accordingly. At the end of the day, everyone is either happy with where they ended up (or who they got), or else they know their spot or their preferred applicant went to an objectively "better match."

As predicted by the prisoner's dilemma, though, the players in this game all choose to play at least a little bit "in bad faith," because both graduate institutions and applicants have superior individual outcomes if they over-represent their quality of their "fit" with one another. So whether it is padding one's resume and rehearsing interview questions on the applicant's side, or a graduate program's "lies of omission" about how difficult it is for graduates to find work after they complete the program, such defections are ubiquitous.

All the players attempt to thwart one another's attempted defection in an ongoing "arms race" of deception and intelligence-gathering, which is based on each players ability to convincingly perform their chosen role, and her ability to "see through" her opponents attempts to present themselves as other than they actually are. In game theory, a "screening move" (Dixit & Nalebuff, 2009) is a move which functions to determine whether a player's behavior is a *sign* (accurate representation) or a *signal* (communicated misrepresentation) of the characteristic under consideration. In other words, a screening move is a strategy I devise to get you to tell me what you're really like, or to determine if you are really like you say you are.

Interview questions are generally "moves" of this kind, for example, in that they simultaneously ask an interviewee to describe their own characteristics but also to demonstrate them in the cognitive task of answering the question, under the interviewer's observation. The applicant's "screening behavior" is likely more subtle, like floating a personal disclosure or a

joke judged to be on the edge of acceptability for the present context in order to "see how it lands," and thus discern his opponent's type. All participants are likely to ask "screening questions" without obviously "right" answers because opponents that are not given obvious signals about how to behave are more likely to behave according to their "true" psychological organization. So when "less prestigious" institutions focus solely on their "welcoming learning environment," applicants may begin to ask about graduates' average income. And when applicants begin to notice that stressing their commitment to multicultural competence "plays well" most places (or doesn't, for applicants of color) and "perform" accordingly, interviewers develop more incisive follow-up questions.

While this could equally be said about personnel selection for a tech startup or graduate students in veterinary medicine, in clinical psychology, personal characteristics play a unique role in the question of suitability for the position. Ivey and Partington (2014 p. 166) note, "Selectors are…tasked with evaluating applicants' reasons for wanting to become psychologists, their psychological stability, emotional maturity and capacity for empathy and self-insight." Though this study focused specifically on how assessors evaluated essays, much of the graduate institution's task in applicant selection amounts to various attempts to measure these kinds of attributes.

In other fields, the limitations of self-report as an assessment method are sidestepped with something like an "audition" or "portfolio." Musicians have long understood that when the band is looking for a new guitarist, it is more efficient to have an "applicants" play the guitar in an "interview" than to ask them to talk about their guitar-playing abilities. This screening move circumvents the tendency of participants in prisoners dilemma of guitarist selection to choose the dominant strategy of "defection" and signal abilities in excess of reality. In psychology, some

programs take a similarly direct tack to the "audition," handing out vignettes which model hypothetical clinical situations, and asking interviewees to provide diagnoses, conceptualizations, or treatment plans; or enlisting their most thespian staff member to be assessed and "case-presentationed" upon by perspiring interviewees. Perhaps more commonly, interviewers perform "unstructured" assessment of factors like "intuition," "self-awareness," or "psychological mindedness" through what they observe, intuit, or infer while the interviewee responds to a series of more-or-less "standard" interview questions.

In their 2014 study, Ivey and Partington studied a group of admissions essay readers at a clinical psychology graduate program assessed and selected for a characteristic they called "woundedness" which they noted "include[s] experiences of early object loss, loneliness, emotional deprivation and absence of appropriate intimacy, failure of carers to meet developmental narcissistic needs, and guilt about not having lived up to parental expectations (p. 167). They found that "the preferred applicant protocols chosen by almost all participant selectors (nine out of ten) as being most suitable for clinical training evidenced woundedness, and almost all participants considered protocols where this was absent as least suitable" (p. 170). The selectors interviewed made a distinction between woundedness and "impairment" which they said designated "problematic psychological states implicated in deficient professional attitudes and behavior" (p. 168). Within this one "factor" where readers are subjectively assessing that enough woundedness is present without rising to the excess of "impairment," a strategic choice is being made to select for a particular range of a particular construct, while the assessors have simultaneously developed a "technology" (whether formalized or just commonly understood) for measuring it.

The assessment of this characteristic is performed in a strategic space with another interactive participant, the applicant. Whatever the applicant's conscious or unconscious motivations in the interaction, the interviewer assesses this and other characteristics attempting to take the interviewee's strategic behavior, his "screening" moves, and his emission of signals rather than signs, into account. The iconic scene from *The Princess Bride* (1987) provides a useful model here:

> Now, a clever man would put the poison into his own goblet because he would know that only a great fool would reach for what he was given. I am not a great fool, so I can clearly not choose the wine in front of you. But you must have known I was not a great fool—you would have counted on it—so I can clearly not choose the wine in front of me.

If the interviewee can intuit that an interviewer might be looking for that sweet spot of woundedness (and since she's made it to the interview, let's give her intuition the benefit of the doubt), she would be clever to say, "I am exactly the right amount of wounded," regardless of her true type. Because the interviewer is not a great fool, he knows that he could be observing a signal rather than a sign, and so he makes the strategic choice to prompt her to *demonstrate* it. But neither is she a great fool, and so anticipating this, she had surreptitiously consulted with successful applicants and hammered out answers to the common interview questions she found on the internet. And so he chooses to "weight" answers based on how spontaneous they sound, and she learns to make her rehearsed answers *sound* spontaneous, and so on.

Another variation on screening moves might be employed by an applicant or program that judges themselves likely to be desirable in the pool and that has strong preferences in terms of their potential "matches." Such an applicant or program might elect to be more honest about their controversial attributes, or if they are especially calculating or audacious, say something

that *only* their favorite kind of match would appreciate. If successful, this move functions to determine and select for the chosen type of applicant or program, though they are banking on the right kind of match being able to recognize this signal. A graduate program or training site that plays up the heavy demands placed on students employs a similar tactic in prompting distress-tolerant, talented, or grandiose applicants to self-select. The same move may also serve other functions such as being an "honest" warning, or a means for faculty and alumni to buttress personal pride about the "elite" nature of the organization with which they identify.

The Strategic Implications of Learning, Nash Equilibria

In the repeatedly aforementioned Prisoner's Dilemma, it is generally possible for the players to cooperate if the game is repeated over and over because, "The successive rounds of the game give [the players] the opportunity to build up trust or mistrust, to reciprocate or placate, forgive or avenge. In an indefinitely long game, the important point is that we can both win at the expense of the banker rather than at the expense of one another" (Dawkins, 1976).

If I know that right after this round of "the prisoners dilemma" we are going to play another one, and then another one after that, I start to notice that what I do in this round might influence your behavior in subsequent rounds. I might start to say things like, "If you cooperate with me this round, I will cooperate with you next round," and you suddenly have the opportunity to observe what I do this round, and use that information to decide whether or not to believe my promise next round. When the formal iterated prisoner's dilemma is played in practice, one of the strongest strategies is known as "Tit for Tat" which "begins by cooperating on the first move and thereafter simply copies the previous move of the other player" (Dawkins, 1976 p. 271). When Tit for Tat plays against itself, it has a totally cooperative series of games, but when it plays against strategies which defect without provocation, the pair fares worse overall. When two

strategies that defect without provocation play against one another, they do still worse. (Dawkins, 1976)

Significantly, this is a game in which "play" consists of decisions between *only two* mutually-exclusive options; and yet despite its simplicity, it is immediately tempting to start to draw comparisons to nuanced aspects of human psychology as "personality traits" (McCrae & Costa, 2010). The comparison in the fates of "Tit for Tat" and more avaricious strategies in this experiment certainly have at least superficial similarity to a conclusion that a therapist might be tempted to draw about the relative "adaptiveness" of a client's tendency to cooperate rather than to defect against trusting others. When it comes to human affairs, we just cannot quite decide whether "what goes around comes around," or "no good deed goes unpunished," but different contexts seem to favor one or the other view.

In the iterated prisoner's dilemma, *because it is repeated,* cooperation triumphs over self-interest. The pre-built strategies Dawkins describes developed relationships with one another, and based their behaviors on idiosyncratic memory of those relationships, to a level of complexity equal to the complexity of the game (the relatively simple choice between *cooperate* and *defect*). The exact opposite is true of the one-off version of the game, in which defection is the dominant strategy. These "strategies" are single instances of the class of entities Harari (2017) describes as "algorithms" (a concept which can be extended to all living organisms, if one views them as "data processors"). In relationships between people in human society, clients and therapists, supervisors and supervisees, students and professors, applicants and interviewers, analogous systems of behavioral predictions based on extrapolation from learning history guide the players' strategic decision-making, dependent on the players' perception of the game they are playing.

The mathematician and game theorist, John Nash, famously the subject of *A Beautiful Mind* (1998), introduced a concept that later became known as a "Nash Equilibrium," which "is a system of beliefs and a profile of actions for which each player is playing a best response to his beliefs. . . a profile of strategies for which *each player* is choosing a best response to the strategies of *all other* players" (Tadelis, 2013 p. 80). Significantly, when all players are playing Nash equilibrium strategies, their "best play" in pursuit of their individual agenda is to continue playing the same strategy, assuming no one else changes tack (a safe assumption if the other players are also rational).

If we assume that any human behavior is an individual's perceived optimal response to any given game, it constitutes a Nash equilibrium between some combination of games played simultaneously at intrapsychic, interpersonal, intersubjective, systemic, and cultural levels. As such, the equilibrium strategy (which in the "algorithmic" view of organisms equates to an algorithm's moment-to-moment "output" in response to input-stimuli) is the set of behavioral "rules" that the individual has adopted as the best response to their developmental environment. "The capacity to interpret human behavior—to make sense of each other—requires the intentional stance: 'treating the object whose behavior you want to predict as a rational agent with beliefs and desires' (Dennett, 1987 in Fonagy et al, 2003 p. 416-417). At the same time, individuals' predictive calculations in intersubjective interaction necessarily utilize idiosyncratic organizing principles, meaning that no two individuals will "output" the same "predictions" (or subjective understandings), even when placed in a superficially identical situation. A client making sense of a therapist's behavior relies on observations of the therapist's behavior in conjunction with pre-existing understanding of what similar behavior has "meant" before. As such, the client's decision to employ a *novel* strategy in response to the therapist may depend on

the therapist defying the client's expectations to a degree that the therapist's behavior does not *make sense* within the client's existing organization.

Commitment Moves, "The Frame," Psychotherapy, and Supervision

In game theory, "commitment moves" are actions a player may take with the intention or the effect of influencing which moves the other player(s) make(s). This often involves communication, as "when one player can move first and make his move known to the other player, or when the players can talk to each other before they move" (Straffin, 1993 p. 85). A simple example of a commitment move would be telling an adversary, "If you hit me, I will hit you back." The promise of a certain behavior (hitting back) is designed to dis-incentivize the opponent from taking an undesirable action (hitting in the first place) and seeks to alter the opponent's choice of strategy. The effectiveness of a commitment move is dependent on its *credibility*, meaning that my promise to behave in a certain way given a certain condition only affects my opponent's decision-making to the degree to which she believes that *my having made the promise* will affect or reflects upon my subsequent decision-making. In other words, my friends might be more likely to lend me money than a stranger would be, partially because my friend is more likely to perceive my promise to pay her back as credible (unless I have a particularly large outstanding debt with that friend already or with people she knows about).

Another example would be the commitment in therapy and supervision relationships—made explicitly or implicitly by the therapist or supervisor—that "under no circumstances will we ever have sex." Relevant APA ethics codes include:

> 3.02 "Sexual Harassment," 3.05 "Multiple Relationships," 3.06 "Conflicts of Interest," 7.07 "Sexual Relationships With Students and Supervisees," 10.05 "Sexual Intimacies With Current Therapy Clients/Patients," 10.06 "Sexual Intimacies With Relatives or

Significant Others of Current Therapy Clients/Patients," 10.07 "Therapy With Former Sexual Partners," and 10.08 "Sexual Intimacies With Former Therapy Clients/Patients" (American Psychological Association, 2017)

Though the therapist or supervisor may conceptualize this commitment as an "ethical principle," it is also a commitment move in that one of its functions is to affect the strategic decision making of the client or supervisee. To the degree to which the client or supervisee perceives this commitment as *credible*, they mentally remove possible outcomes which involve a sexual relationship between themselves and their supervisor/therapist from their "psychological payout table" and alter their behavior accordingly. In practice, were this not the case, a hypothetical client who finds his therapist attractive might be well advised to selectively disclose only those thoughts and recollections which he thinks would be most likely to incite his therapist's interest. When the therapist communicates a credible commitment to abstain from this outcome *no matter what*, the client's strategic calculus shifts, potentially toward the slightly less attractive "second place prize" of psychological growth, and disclosure of thoughts and recollections deemed most aligned with its service. Unsurprisingly, the incentive structure that emerges when the therapist feels free to initiate a sexual relationship with their attractive client and act accordingly in therapy is even more markedly opposed to the agenda of therapeutic change.

Graduate students of clinical psychology sit in a crossroads between arguably the most nuanced and ambiguous strategic situations encountered in a typical clinical career: providing therapy, being supervised, and supervising. To varying degrees depending on the personal strategies of the participants in any of these activities, the ambiguity of these situations are mitigated by a set of commitment moves wrapped up in the widely-discussed construct of the

"frame" of psychotherapy. (Langs, 1978) This construct incorporates both explicitly discussed "rules" (confidentiality and its limits, handling of dual relationships, etc.) and implicitly communicated "stances" which are gradually manifested through iterated interaction.

The therapist's prompt beginnings and endings, if they are consistent, serve as a credible commitment to continue in to begin and end sessions on time in the future. This may have only practical effects on the dependability of clients' and therapists' schedules, or it may create for the pair a profound sense of security in the appointed hour; and for the client, grant the strategic freedom to enter into intimidating experiential states secure in the knowledge that "it will be over at 4:50, for sure." In intersubjective theory, the therapist's emotional response to the client's articulated and felt experience is seen to organize and reorganize that experience. In the playspace of the intersubjective field, the psychotherapist chooses her moves with the aim of creating opportunities for adaptive reorganization. In evaluating the strategic alternatives of various possible "things to say," she aims at "attunement" and looks for utterances she predicts will both resonate with—and help "make sense of"—the client's subjective experience (Buirski, 2005). She uses her own lifetime of experience as the working data pool that constitutes the predictive model she uses to calculate likelihood that a particular utterance will be received as "attuned," referenced against her collected understanding of the client's intersubjective organization and her particular position within it.

In supervision, whether vocalized or communicated implicitly through consistent behavior, supervisors may use commitment moves in an attempt to create goal-compatible incentives for supervisees. The supervisor holds the dual roles of "evaluator" and "training provider" in the game of supervision, and these roles come with conflicting agendas. Trainees "work to please their supervisors not only because of the immediate power differential but also

because of potentially lifelong ramifications for their careers, including grades, letters of recommendation, clinical appointments, and so on" (Karson, 2008 p. 201). When the evaluative role goes unmitigated by the supervisor's commitment moves, the clever student chooses to show only their best tape, and summarize sessions in a way designed to highlight their competence, rather than to foster its development. A supervisor's decision to adopt a policy of categorically abstaining from "praise," or responding with neutral curiosity to good tape and bad tape alike—both work to balance the emotional reward of discussing the things the supervisee actually needs help with.

Consider the differences in strategy logically employed by a student whose primary agenda is "feeling like he is a 'good' therapist" compared to one rewarded by "moments where she notices and makes use of an opportunity for improvement." In the first case, the student's equilibrium strategy will likely be something like finding the lowest-effort route to a standard subjectively perceived as just enough to meet the "good" benchmark, conceal his flaws from himself and others, and do her best to reject novel information that could threaten his "good-therapist" identity. The therapist who gets off on improvement, meanwhile, rationally chooses to seek and reimagine flaws in her technique, and to effortfully pursue novel and potentially superior methods. Though the supervisor will encounter students that emit a variety of signals about their subscription to one or the other agenda, he is also able to play commitment moves which favor one agenda over the other.

Mechanism Design, The Match

A subfield of Game Theory called "Mechanism Design" inverts game theory's strategic consideration by starting with a strategy the "designer" wants the players to play, and designing a game in which rational agents will behave as the designer intended (Börgers, Krähmer, &

Strausz, 2015). In a sense, a mechanism is a strategic analog of an operant conditioning chamber or "Skinner Box" in that an agent playing a "designed mechanism" can be induced (through rational self-interest) to play a predetermined, "intended" strategy (pushing the "lever") if structure of the game is designed so that "pushing the lever" is the optimal strategy.

One of the most common of these "intentions" in mechanism design is known as "preference revelation" (McMillan, 1979) which describes strategic contexts in which players who choose their optimal or "equilibrium" strategy, in so doing, reveal their true characteristics. In practical terms, this is often the intention of auction houses trying to design a system that induces buyers to pay the highest amount they would be willing to pay for each item for sale.

An "English Auction" is where an "auctioneer stands at the front of the room calling out ever-increasing bids" until only one bidder remains willing to pay the current bid. (Dixit & Nalebuff, 2008, p. 303). This is an example of a game for which the optimal strategy is for players to bid their true highest value for goods, and thus reveal their "true" personal valuation preferences. Preference revelation is also of central importance in the Association of Psychology Postdoctoral and Internship Centers (APPIC) national internship match where a "matching algorithm" presides over the process of funneling prospective interns to internship sites. The task is modelled in a strategic construct called the "Stable Marriage Problem" in which agents of two types (here, "interns" and "internships") are tasked with "pairing up" with mutual consent. The agents in the system are assumed to have the agenda of picking the most attractive "partner" according to their own personal taste. The "question" in the stable marriage is how to achieve a situation where no one has anything to gain by leaving their partner for someone else, and all the "marriages" are therefore "stable" (Dubbins, 1981). If this condition were met, it would also

mean that the principle of preference revelation had been satisfied as all parties had behaved "honestly" regarding their preferences about partners.

For all practical purposes, the "answer" to this question is "The Gale-Shipley Algorithm" (Dubbins, 1981) which in a slightly modified form, constitutes the core computational process of the APPIC Internship Match. Very roughly, it performs a series of iterations in which interns are all paired up with their highest remaining preference in available internships, and then the stragglers are allowed to "bump off" anyone to whom their prospective next-most-preferred internship prefers them. Because this is modelled virtually, the computer running the algorithm can perform all possible pairings simultaneously and end up with only "stable" pairings, and "revealed" preferences (National Matching Services, 2017).

What makes this algorithm so significant in the match (and tasks like it) is that neither internships nor interns can fare better individually by deceptively reporting their preferences about one another than they would by reporting preferences honestly. And so they arrive at something much closer to outcome of the utopian vision of "honest" graduate applicant selection presented earlier than would be the case in the more defection-rewarding situation in which interns are "hired" on a one-by-one basis, and each pairing must be either taken or left *permanently* before other offers are considered.

Though the task of actually assessing partners is performed in the same strategic fog as described previously in graduate admission interviews, the quality of "preference revelation" built into the ruleset of the match itself means that the applicant's question, "Should I take this one? I have no idea if I'll even get another offer," and the myriad strategies that immediately erupt around it, are rendered obsolete. This is a system built *with the intention that* "honesty pays," and within its domain of influence, it enforces a "dominant strategy" of cooperation.

Tasks of mechanism design appear on multiple levels in the context of graduate training, from a therapist's particular style of frame-setting, to a program director's decisions about organizational policy. These designers generally want the game for which they are writing the "rulebook" to have certain characteristics, though the desired characteristics vary widely based on the situation.

Intra-systemic Conflict, Program Design

Though graduate institutions and training sites may set out to "create good psychologists," various factors may subvert this goal in various ways. The dictates of "emotional safety" pathologize truth-telling, and the assumption that "we're all just here to help people" blinds us to our own psychology. Organizational cultures conspire to expel potentially necessary new ideas, and program administrators select for evidence that "all is well" just as cracks appear in the institutional foundation.

In multi-agent systems, even when agents are genuinely seeking to cooperate, there is conflict between alternative cooperative strategies, and even a therapist focused *exclusively* on creating a nurturing "holding environment" inevitably comes into conflict with her client's pre-existing organization of experience. Strategies that systems employ for handling internal conflict are visible in the conflict resolution protocols in graduate programs' policy manuals, and the spectrum of ways that supervisors and therapists respond to their counterparts' challenges. In some organizational cultures, classrooms, therapies, and individual psychologies, the stable strategy is some combination of disavowal, willful ignorance, and mandated self-censorship. This strategy may have a range of "side effects" such as consolidation of power (the dissatisfied must not complain), slow development (it's hard to suggest improvements without disagreeing), and systemic fragility (new information and new participants can be devastating).

Any dynamic system necessarily relies on information-gathering about its own functioning in order to maintain functional equilibrium—a basic tenet in cybernetics (Bateson, 1976). For graduate training institutions, the ability of administrators, instructors, supervisors, and therapists to gather accurate information about the functioning of the systems they oversee is of central importance to the constant, repeated corrections necessary to keep the system on course. Rodriguez-Menendez et al (2017) point out two ways in which a program's self-assessment is likely to be confounded: "Program faculty may be making decisions and forming an understanding of situations based on their own perceptions of the student experience, regardless of whether the information is accurate," and "program administrators and faculty may focus on those aspects of curricular delivery and clinical supervision which confirm the perceived views about the effectiveness of their curriculum, while ignoring aspects of their model which disconfirm their operational hypotheses" (p. 5).

In other words, program administrators may *see* good psychologists being created when this is not the case. Though this problem is to some degree common to any "administration," clinical psychology also has some unique struggles, such as a peculiar tendency to promote a view of its participants as *primarily* altruistic, and to stigmatize self-interest and overt conflict. Ivey and Partington (2012 p. 167) note that "typically, applicants for clinical psychology training claim some variation of the altruistic wish to alleviate people's suffering" as their primary motivator in pursuing psychotherapy as a profession. Institutional cultures which "buy in" to the claims of altruism—seeing it as a sign rather than a signal—design a training system for uncomplicated altruists which places significant constraints on the engagement of other motivations. Such a program (likely staffed by faculty and administrators that make similar claims) addresses motivations like professors' self-aggrandizement, therapists' voyeurism, and

students' desire for leisure—through simple avoidance, and may miss out on the opportunity to confront them productively. As a counterexample, a strategy employed in the University of Denver Graduate School of Professional Psychology's admissions essay prompt, which includes the instruction: "Avoid writing about the wish to help others or about how you want to contribute to society," ("PsyD Essay Questions," 2018) may alternately prompt a demonstration of self-reflective capacity and serve as an implicit commitment move to incentivize honest self-assessment in future play (assuming essay readers enforce it).

"Nature," a somewhat older system than clinical psychology graduate training, provides an interesting case study in a similar principle. In discussing the effects of the disappearance of large predators in a variety of ecosystems, Stolzenberg (2009) found that the presence of an apex predator in a given environment seems to actively preserve the biodiversity of the organisms on multiple tiers of the food chain below it. The predator fosters the diversity of both carnivorous and herbivorous prey animals by preventing any one species gaining dominance, as well as diversity among plant life by teaching grazing animals to avoid the otherwise-tempting dense foliage where ambushes are most easily sprung. Predators defend against scenarios in which a species of herbivore multiplies unrestrainedly, inevitably narrowing the host system's plant life down to only those plants which the grazing animal cannot eat. Conflict in ecosystems can be understood strategically by assuming that every organism's "agenda" is to preserve and promote its particular genetic makeup (Dawkins, 1976).

Similarly, in designing the ruleset for the multi-agent system of the new United States of America, the authors of the Constitution made the strategic decision to assume self-interest on the part of individual members of the constituent populace. The hope was that self-interested individuals could be placed in balanced opposition, rather than attempting to legislate against

"selfishness" and in so doing, brutally curtail "liberty." In *Federalist Paper #10*, Madison articulates his justification for the specific architecture of the government, whose representatives "must be raised to a certain number, in order to guard against the cabals of a few; and that, however large it may be, they must be limited to a certain number, in order to guard against the confusion of a multitude" (Hamilton, Jay & Madison, 2015 p. 45). In his role as a "mechanism designer," Madison proposed to "build in" the power-regulating function that big carnivores serve in ecosystems.

Madison's hope is that both the *nature* of the government—as well as the government itself—will affect the system in which it holds sway in specific ways. First, it becomes more able to accomodate to the introduction of novel strategies, or new "types" of agents. Any specific strategy will be relatively unlikely to both be "killed off" and to "take over." Second, the system seems to become more durable both in a vacuum and in interaction with other systems. Both *mechanisms* are *designed* to handle intra-systemic conflict in ways that shape it to systemic advantage. The presence of the predator species, and the formula Madison offers for governmental representation, both serve to challenge the dominance of coincidentally dominant things within the system.

A therapist and/or client face a similar task to that of the authors of the Constitution as they design the "government" for a society of two. These designers attempt to structure their respective systems in a fashion that promotes things like freedom, collaboration, and the expression of diverse perspectives. A therapist might conceptualize part of what she is doing as helping a client integrate a greater diversity of self-experiences by loosening the grip of whatever particular system of psychological self-oppression he happens to employ. Karson (2008) notes that this aim is ubiquitous among different schools of therapy:

> Every major school of therapy addresses the issue of empowering the marginalized, whether it's Freud's recovery of the repressed, cognitive therapy's disputation of the party line, behaviorism's suspicion about language and rules, systems theory's investigation of oppressive or paradoxical role relationships, humanism's celebration of the tyrannized, or intersubjecticity's care not to blame the patient for the cocreated relationship. (Karson, 2008, p. 118)

The "issue of empowering the marginalized" applies equally in the multi-agent system of individual psychology. If forces within the psyche are like the independent agents in a system, then those agents may best exist in balance."In the psyche, there should be a balance between a desire to satisfy a particular figure, and the needs of the person on the whole… a good psychological organization should also protect the entire populace from tyranny, [and] encourage and facilitate diversity (by recognizing that not everyone is the same, and that as many needs should be met as possible)" (Karson, 2001 p. 64-65). And yet, many of the strategic realities in graduate institutions may function to confound this very goal.

Conclusion

What Stolzenberg prescribes for ecosystems, Madison for governments, and Karson for therapies might be productively applied to the games in which psychotherapists are trained. While the phrase "freedom from tyranny" may sound a bit overwrought for application to things like the protection of critical thinking in graduate school classrooms, the silencing of minority viewpoints robs as much from classrooms as it does from society. Clinical psychology is awash both in the *technology* of productive strategic change, and in sophisticated methods for understanding agentic motivation. The lens of game theory provides a useful system for

understanding how the nature of the games that agents play and perceive affect their strategic play, and viewing training contexts strategically renders them subject to strategic intervention. Game theory encourages a view of suboptimal behavior as an optimal response to a suboptimal game, and in applying this lens to clinical psychology graduate training, it suggests building better games.

This paper has explored a handful of training contexts as "games" and articulated some of the strategic considerations faced by the agents playing them. The exploration has been relatively shallow in scope, and it has almost categorically neglected the issue of what optimal play and optimal game design might actually *be*, in pursuit of creating maximally competent psychotherapists. The determination of optimal play for students, supervisors, professors, and administrators represents a potentially substantial basis for productive future inquiry.

Works Cited

Alper, G. (1993). The theory of games and psychoanalysis. *Journal Of Contemporary Psychotherapy*, *23*(1), 47-60.

American Psychological Association, (2012) Recognition of Psychotherapy Effectiveness. Retrieved from: http://www.apa.org/about/policy/resolution-psychotherapy.aspx

American Psychological Association, (2017) Ethical Principles of Psychologists and Code of Conduct. Retrieved from: http://www.apa.org/ethics/code/

Atwood, G. & Stolorow, R. (2014) *Structures of Subjectivity: Explorations in psychoanalytic phenomenology and contextualism.* New York: Routledge.

Bateson, G. (1972) *Steps to an Ecology of Mind.* Chicago: University of Chicago Press

Börgers, T., Krähmer, D., & Strausz, R. (2015). *An Introduction to the Theory of Mechanism Design*. Oxford: Oxford University Press.

Brandchaft, B., Doctors, S., & Sorter, D. (2010). *Toward an Emancipatory Psychoanalysis: Brandchaft's intersubjective vision*. New York: Routledge.

Buirski, P. & Haglund, P. (2001) *Making Sense Together: The intersubjective approach to psychotherapy.* Lanham: Rowman & Littlefield Publishers

Buirski, P. (2005) *Practicing Intersubjectively.* Lanham: Rowman & Littlefield Publishers

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dixit, A. K., & Nalebuff, B. (1991) *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life.* New York: Norton.

Dixit, A. K., & Nalebuff, B. (2008) *The Art of Strategy: A game theorist's guide to success in business & life*. New York: Norton.

Dubins, L. & Freedman, D. (1981). "Machiavelli and the Gale–Shapley algorithm". *American Mathematical Monthly*. 88 (7): 485–494.

Flabbi, L., & Pediconi, M. G. (2014). Unconscious and Game Theory. *International Journal Of Applied Psychoanalytic Studies*, *11*(4), 339-359.

Fonagy, P., Target, M., Gergley, G., Allen, J. & Bateman, A. W. (2003) The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A theory and some evidence. *Psychoanalytic Inquiry,* 23:3, 412-459.

Freud, A. (1992) *The Ego and the Mechanisms of Defense.* London: Karnac Books

Glazer, B. (Producer) & Howard, R. (Director). (2001) *A Beautiful Mind.* [Motion Picture] United States: Universal Pictures

Hamilton, A., Jay, J. & Madison, M. (2015) *The Federalist Papers: a collection of essays written in favour of the new constitution as agreed upon by the federal convention september 17, 1787.* Coventry House Publishing

Harari, Y. N. (2011) *Sapiens: A brief history of humankind.* New York: Harper Collins.

Harari, Y. N. (2017) *Homo Deus: A brief history of tomorrow.* New York: Harper Collins.

Ivey, G., & Partington, T. (2014). Psychological Woundedness and its Evaluation in Applications for Clinical Psychology Training. *Clinical Psychology & Psychotherapy, 21*(2), 166-177.

Kahneman, D. 2011. *Thinking Fast and Slow.* New York: Farrar Straus & Giroux.

Karson, M. (2008) *Deadly Therapy: lessons in liveliness from theater and performance theory.* Lanham: Rowman & Littlefield Publishers.

Karson, M. (2001) *Patterns of Child Abuse: How dysfunctional transactions are replicated in individuals, families, and the child welfare system.* Lanham: Rowman & Littlefield Publishers.

Kuhn, T. (1962) *The structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Langs, R. J. (1998) *Ground Rules in Psychotherapy and Counselling*, London: Karnac Books.

Lear, N. (Producer) & Reiner, R. (Director). (1987) *The Princess Bride.* [Motion Picture] United States: Act III Communications.

McMillan, J. (1979). *The Free-Rider Problem: A Survey*. Economic Record. **55**: 95–107.

Mitchell, J. (1986) *The Selected Melanie Klein.* New York: The Free Press

Miller, S., Hubble, M. & Duncan, B. (2007) *Supershrinks: What is the secret of their success?* The Psychotherapy Networker. 27-56

National Matching Services, (2017) *The Matching Algorithm.* Retrieved from: https://natmatch.com/aoairp/aboutalg.html

Orange, D. Atwood, G. & Stolorow, R. (1997). *Working Intersubjectively : Contextualism in psychoanalytic practice.* Hillsdale: The Analytic Press.

Rodriguez-Menendez, G., Dempsey, J. P., Albizu, T., Power, S., & Campbell-Wilkerson, M. (2017). Faculty and Student Perceptions of Clinical Training Experiences in Professional Psychology. *Training And Education In Professional Psychology*, *11*(1), 1-9.

Shoham, Y., & Leyton-Brown, K. (2009). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge: Cambridge University Press.

Skinner, B. F. (1965) *Science and Human Behavior.* New York: The Free Press.

Stolzenburg, W. (2008) *Where the Wild Things Were: Life, death, and ecological wreckage in a land of vanishing predators.* New York: Bloomsbury

Straffin, P. D. (2006). *Game theory and strategy*. Washington: Mathematical Association of America.

Tadelis, S. (2013). *Game theory: An introduction*. Princeton: Princeton University Press.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the adaptive unconscious*. Cambridge: Harvard University Press.

Wampold, B. & Brown. G. (2005) Estimating Variability in Outcomes Attributable to Therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, 73(5), 914-23.