Games and Economic Behavior **32,** 220–246 (2000)
doi:10.1006/game.2000.0796, available online at http://www.idealibrary.com on IDEAL®

Variable-Frame Level-*n* Theory[1]

Michael Bacharach

Institute of Economics and Statistics, University of Oxford, Oxford, United Kingdom

and

Dale O. Stahl

Department of Economics, University of Texas at Austin, Austin, Texas

Received December 10, 1997

We develop a boundedly rational version of variable frame theory by merging the variable-frame concept with level-*n* theory. Variable frame theory assumes that a player's options are determined by a set of attributes (her "frame"), which induce a partition of the action set. Schelling competence emerges by combining frame-induced options with unbounded rationality and with payoff-dominance and symmetry disqualification. The weak empirical evidence for these assumptions motivates variable frame level-*n* theory, which predicts the Schelling competence by combining a nonrational level-0 tendency with reasoning by higher-level players. *Journal of Economic Literature* Classification Numbers: B41, C70. © 2000 Academic Press

1. INTRODUCTION

1. *Rational Variable Frame Theory* (*VFT*)

Variable frame theory (VFT) was introduced in Bacharach (1993) as an alternative foundation for and as a refinement of Nash equilibria. Bacharach and Bernasconi (1997) applied VFT to experimental pure coordination games and found considerable support for the general approach, but at this point in time the evidence for the ancillary rationality assumptions remains mixed. The purpose of this paper is to develop a boundedly

[1] This paper was written as part of the research program "Framing, Salience, and Product Images" under ESRC Contract L122251002.

0899-8256/00 $35.00

rational version of VFT by merging the variable-frame conceptual apparatus with the level-*n* theory (LNT) of Stahl and Wilson (1995).

We shall confine ourselves to object-choosing games, in which each player must choose one object from a finite set and payoffs depend solely on which objects are chosen. This class of games is very wide: for example, we may consider different monetary amounts demanded in a bargaining game as different objects. Indeed, any game with a common strategy set can be formally regarded as an object-choosing game.

A particularly simple class of object-choosing games are *matching games* (Bacharach, 1993; Sugden, 1995; Janssen, 1997). A matching game is an object-choosing game in which, for each $a \in A$, each player receives the amount $y_a > 0$ if all choose a and receives nothing otherwise. A matching game is called *plain* if y_a is constant for all a.

A *frame* on A is a collection of families of properties of the objects of A that can be used to agglomerate them into groups of like kind. Given a set A, frames on A may vary in discriminating power. For example, if A is a set of physical objects which differ in terms of color and shape, then there are several possible frames: both color and shape families might be used to discriminate the objects, or just color, or just shape, or neither. If A consists of a red cube and three balls, one blue, one yellow, and one red, then the shape frame, the color frame, and the composite shape/color frame partition A as {*cube*, *balls*}, as {*red*, *blue*, *yellow*}, and as {*red cube*, *blue ball*, *yellow ball*, *red ball*} respectively. Finally, we define the "empty frame," which discriminates none of the objects from each other and induces the trivial partition $\{A\}$. We will call the elements of a partition "cells." A frame made up of a subset of the families of a given frame we will call a "subframe" of the latter: for instance, the color/size frame is a subframe of the color/shape/size frame.

A fundamental feature of VFT is that a player's options are determined by her frame—specifically, by the cells of the partitions of A induced by her frame and its subframes. The underlying principle is that to choose an object a player must have some way of "mentally fixing" that object. This is usually a definite description, such as "the yellow ball."[2] If this were the only possibility, then if a player had the shape frame in our four-object example there would be only one object she could choose: the cube. However, VFT assumes that her ability to categorize the balls as balls gives her a second option, which consists in choosing at random, or "picking," one of the three balls. Similarly, if she has the empty frame and only sees the objects of A as objects, she has the option of picking any object. Thus, to each cell of a frame-induced partition is associated one "option."

[2]But it could be a "nonconceptual representation" such as a certain ineffable look or smell.

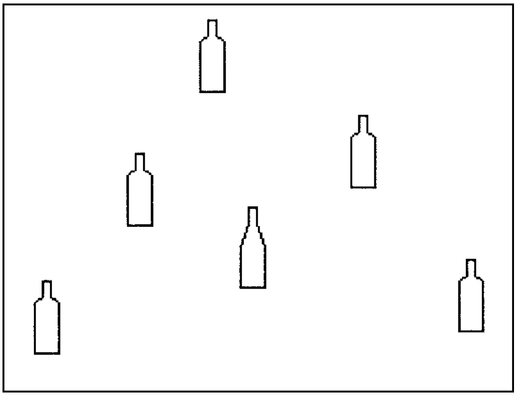

FIG. 1. Bottles.

Example 1: *Bottles*, a "Simple Schelling Game".[3] There are K bottles ($K \geq 3$), all identical except that one is hock-shaped and the rest are claret-shaped (Fig. 1). Two players indicate their choice for an object simultaneously and privately. If both players choose the same bottle then each receives a payoff of 1, and otherwise they get nothing. The shape frame partitions the bottles into two cells (*hock-shaped*, *claret-shaped*), and hence induces two options: choosing the hock bottle (h) and picking a claret bottle (c). The empty frame (a subframe of the shape frame) induces a third option of picking any bottle (b). Thus for a player having the shape frame the option-induced payoff function is

	h	c	b
h	1	0	$1/K$
c	0	$1/(K-1)$	$1/K$
b	$1/K$	$1/K$	$1/K$

There are three pure-strategy Nash equilibria in the options: hh, cc, and bb; in addition, there is a continuum of mixed-strategy Nash equilibria, but these are not relevant to rational VFT. Payoff dominance (PD), an auxiliary axiom of rational VFT (Bacharach, 1993), selects the hh equilibrium, which is the Schelling solution.

[3]A "simple Schelling game" (Bacharach and Bernasconi, 1997) is a plain matching game in which there is just one family of properties for classifying the objects, and this family defines a single "oddity" (an object which is the only one of its kind). Later we shall also see "discrete Schelling games," in which there is a second family of properties which partitions the objects discretely, and "hi–lo Schelling games," in which the payoff for matching varies over objects.

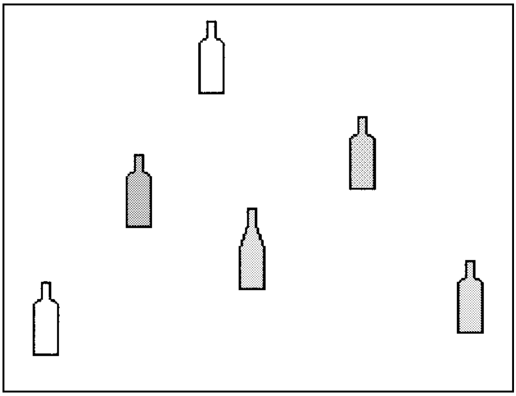

FIG. 2. Colored bottles.

EXAMPLE 2: *Colored Bottles*, a "Discrete Schelling game." There are K bottles ($K \geq 3$), each a different color, one hock-shaped, the rest claret-shaped, and otherwise identical (Fig. 2).

In this example, color (if noticed) generates an enriched frame, and this color/shape frame induces the discrete partition of the objects and hence K additional options of the form "choose the ...-*colored bottle*." (Since the shape frame is a subframe of the color/shape frame, the player still has the shape-induced options.) We now have K additional pure-strategy Nash equilibria, all of which give an expected payoff of 1, like hh. To resolve this multiplicity, Bacharach (1993), Janssen (1997), Bacharach and Bernasconi (1997), and Casajus (2000) invoke the principle of symmetry disqualification (SD), which states that if two cells of a frame-induced partition contain the same number of objects then, since neither can be rationally presumed to be more probable than the other, any Nash equilibrium which contains the options given by these cells is disqualified as a solution to the game. In Example 2, all the cells of the color-induced partition contain exactly one object, and so all the K pure-strategy Nash equilibria in which each player chooses the same color-induced option are symmetry-disqualified as solutions.[4]

What, the reader might ask, if one bottle is reddish and the others are all greenish but of slightly different shades? Would (SD) still be applied? The answer is "Yes" with respect to the color-induced discrete partition. However, now another frame is possible, namely {*reddish*, *greenish*}. This frame would lead to (*reddish*, *reddish*) as a Nash equilibrium, and this equilibrium would not be disqualified.

[4]It is however possible for players to choose the hock bottle *qua* hock-shaped, as we shall see.

The first moral of the variable frame approach is "frames matter"; the second is "beliefs about frames matter." What makes it sensible for Player 1 to choose an object with a certain property, say being plastic rather than glass, instead of choosing, say, the only hock bottle, is having reason to think that Player 2 will choose the plastic bottle; but Player 2 can make this choice, under this description, only if the plastic–glass classifier is in his frame. Rational players must therefore have beliefs about the other players' frames. Rational VFT imposes a restriction on these beliefs: the no refinement principle (NR) that the belief spaces considered possible for another player cannot refine a player's own belief space. For example, if a player is not aware of any shape differences in Example 1, then she cannot have a belief about the other players' shape-induced options; for if she has such a belief *ipso facto* she *is* aware of shape differences.

Not only does VFT resolve game-theoretic indeterminacy in the above games, but the solution it defines also explains the long puzzling "Schelling competence." This is the ability to solve coordination problems such as rendezvous by collectively alighting on a salient equilibrium or "focal point" (Schelling, 1960; Mehta, Starmer, and Sugden, 1994). In the traditional analysis of Example 1, there is one equilibrium for each bottle (in which both players choose that bottle). Intuitively, one of these, in which both choose the hock bottle, stands out or is *salient*, at least for players who notice the shapes; traditional theory fails to predict that shape noticers will choose the hock bottle, but VFT does. In Example 2, going for the hock bottle also seems, intuitively, salient; traditional theory offers no prediction while VFT predicts that shape noticers choose it.

2. *Derationalizing VFT*

The resolving power of VFT in Examples 1 and 2 depends heavily on the refinement principles (PD) and (SD). These are putative principles of rational behavior whose basis is not secure. Although (PD) is very reasonable in pure coordination games, there is evidence from mixed-motive games (e.g. Van Huyck, Battalio, and Beil, 1990) that when pitted against risk dominance it often comes out on the bottom, and the literature as yet contains no compelling *a priori* arguments for this. (SD) is also plausible, but it has not yet received serious empirical support in "hard cases." Moreover, the Nash equilibrium principle itself, which they refine, is an *a priori* principle of rational choice with mixed theoretical and empirical support (Stahl and Wilson 1995).

The unconfirmed status of (PD), (SD), and the Nash equilibrium principle motivates what follows, in which we seek to apply the basic ideas of VFT to the boundedly rational level-n theory (LNT) of (Stahl and Wilson, 1995). We shall call the resulting theory the "variable frame level-n theory"

(VFLNT). It contains neither (PD) nor (SD), and in it Nash equilibrium is a principle obeyed by only one specific type of player among several.

It might be wondered how much rationality is involved in the variable-frame element of VFT, which we retain. The answer is, we think, only an irreducible minimum. Every player, however limited her reasoning powers, is a concept possessor; otherwise, she could not understand her situation at all. And as long as players are concept possessors, and there are alternative concepts which they might use in representing a particular choice situation, the variable frame element of VFT necessarily applies.

True, there is *scope* for more rationality than this in framing a decision problem. In particular, it may be that agents engage in an optimized search activity in their framing, actively seeking richer or more convenient representations (Newell and Simon, 1972). This is a research direction in variable frame theory which we pursue elsewhere.

3. *Layout*

In Section 2 we summarize rational VFT, whose formalism we will take over for VFLNT. In Section 3 we present the VFLNT. We first describe the main features of level-0 players, then those of players of levels 1 and beyond. Next, we derive a general characterization of play at all levels by players having these features. We illustrate our formal discussion throughout with two running examples (Bottles and Colored Bottles). These are coordination games of the kinds studied in (Schelling, 1960; Bacharach, 1993; Sugden, 1995; Janssen, 1997; Bacharach and Bernasconi, 1997; and Casajus, 2000): it is important to consider such games, for it is a critical test of the bounded rationality version of VFT that it should share the predictive success of the full rationality version for such games presented in Bacharach (1993) and Bacharach and Bernasconi (1997). In Section 4, we apply the general characterization of play to specify level-1 and level-2 play for our running examples, and we prove an exact condition for the VFLNT to yield coordination on a focal point in a wide class of Schelling-like games. In the final section we discuss testable differences between VFLNT and VFT, assess what we have learned, and raise some open questions.

2. A REVIEW OF VFT

Formally, we have a finite set A of *objects* and a set $\mathscr{C}$ of *classifiers* of A. Each element c of $\mathscr{C}$ is a finite set of properties which classify the objects of A exhaustively and mutually exclusively. Thus each classifier c induces a partition, P_c, of the object set A. For example, if A is a set of physical objects, typical classifiers of A are a set of colors, a set of shapes, and so on.[5] The cell of P_c corresponding to a particular property of c is called the *extension* of that property; e.g., the subset of the objects of A which are blue is the extension of *blue* in A.

A *frame* is a subset of $\mathscr{C}$, that is, a set of classifiers. Let $\mathscr{F}$ denote the set of frames, i.e., $\mathscr{F} = \mathscr{P}(\mathscr{C})$, the power set of $\mathscr{C}$. We will use sans serif font to indicate frames. Each frame $\mathsf{f} \in \mathscr{F}$ provides a cross-classification of the objects of A, associated with a set of (in general) conjunctive properties, and hence each frame f defines a partition P_f of A. The cell of P_f corresponding to a particular conjunctive property is called the *extension* of that property. For example, a frame consisting of a shape classifier {*round*, *square*, *triangular*} and a color classifier {*blue*, *green*} cross classifies the objects by conjunctive properties such as *square* $\wedge$ *blue*, and *triangular* $\wedge$ *green*. The set of objects of A that are square and blue is the extension of *square* $\wedge$*blue*. The empty frame represents the possibility that a player cannot sort the objects of A in *any* way, but can only gather them all under some embracing description such as "the objects" or "these bottles."

We take the distribution of frames in the population of players to be exogenous and let the probability measure $v(*)$ represent that distribution. $v(\mathsf{f})$ may be thought of as the *salience* of the frame f. Bacharach and Bernasconi (1997) discuss some of the cultural and contextual determinants of $v(*)$.

A player's *options* are defined as follows. To each (in general, conjunctive) property o of f is associated an option (also denoted by o) and its extension $E(o)$.[6] If $E(o)$ is a single object, then "to choose the option o" implies choosing that object. If $E(o)$ is a nonsingleton set of objects, then "to choose the option o" implies *picking* an object in $E(o)$ (randomly with equal probability). In the latter case, we shall say that the player *realizes* the object selected by the random process. We let opt_f denote the set of options so defined by the frame f.[7]

If f is a frame in $\mathscr{F}$, write $F(\mathsf{f})$ for the set of all subframes of f. Let $\mathrm{opt}_{F(\mathsf{f})}$ denote the complete set of options defined by all the subframes of f, that is, $\mathrm{opt}_{F(\mathsf{f})} = \bigcup_{\mathsf{f}' \in F(\mathsf{f})} \mathrm{opt}_{\mathsf{f}'}$. Finally, let $\mathrm{Opt} = \bigcup_{\mathsf{f} \in \mathscr{F}} \mathrm{opt}_{F(\mathsf{f})}$. Let a player's

[5]What we here call classifiers are called *families* in Bacharach (1993) and Bacharach and Bernasconi (1997), *dimensions* in Janssen (1997), and *attributes* in Casajus (2000). One can think of a classifier simply as a *variable* which takes a finite set of values, but in the present context the variables are typically physical and other everyday properties by which people sort objects, and the terminology "families of properties" or "classifiers" is more intuitive.

[6]We use the phrases "extension of property o" and "extension of option o" synonymously and denote both by $E(o)$. Note that a subset of objects may be the extension of more than one property, belonging to different frames, and thus the extension of more than one frame-conditioned option. This is true, for example, of the hock bottle in Colored Bottles.

[7]Note that the set of extensions of the options in opt_f is the partition P_f of A which is the join of P_c over $c \in \mathsf{f}$.

payoff be given by the function u from $A \times A$ into the reals. Then it is straightforward to derive a payoff function U which maps from Opt $\times$ Opt into the reals as

$$U(o, o') = \sum_{a \in E(o)} \sum_{a' \in E(o')} u(a, a')|E(o)|^{-1}|E(o')|^{-1}, \tag{1}$$

where $|*|$ denotes the cardinality of a set.

A *decision rule* in VFT is a function, ϕ, which maps each possible frame f into an element of $\text{opt}_{F(\mathsf{f})}$, the set of options given by that frame and all its subframes.

In addition to a frame, a rational VFT player must have beliefs about the frame of her coplayer. Let $\hat{v}(\mathsf{f}' \mid \mathsf{f})$ represent the subjective probability for a VFT player with frame f that her coplayer has frame f'. The (NR) principle prohibits a player from putting any positive weight on any frame that is not a subframe of her own frame; that is,

(NR) The support of $\hat{v}(* \mid \mathsf{f})$ is a member of $F(\mathsf{f})$.

The (NR) principle is, as we have already pointed out, a necessary implication of the very meaning of a frame. Bayesian rationality requires that $\hat{v}(* \mid \mathsf{f})$ be formed from the true exogenous distribution v, by truncating v to $F(\mathsf{f})$ and rescaling.

Given a decision rule ϕ for the coplayer and subjective probabilities for the coplayer's frame conditional on the player's frame f, the expected utility of each option in opt_{f} is well-defined. Specifically, for any option $o \in \text{opt}_{F(\mathsf{f})}$,

$$\mathscr{E}U(o \mid \phi, \mathsf{f}) = \sum_{\mathsf{f}' \in F(\mathsf{f})} U\big(o, \phi(\mathsf{f}')\big)\hat{v}(\mathsf{f}' \mid \mathsf{f}).$$

Then (ϕ, ϕ) is a *variable frame equilibrium* (vfe) if, for each f, $\phi(\mathsf{f})$ maximizes the player's expected utility against $\phi(*)$.

VFT further imposes the two following principles, payoff dominance and symmetry disqualification:

(PD) If a vfe is strictly payoff dominated by another vfe, then the payoff-dominated vfe is disqualified as a solution.

(SD) If two cells of a frame-induced partition contain the same number of objects, any vfe which does not assign equal probability to these cells is disqualified as a solution.

A vfe that satisfies (PD) and (SD) is called a "variable frame solution." Despite these additional refinement principles, a game could have multiple variable frame solutions, though matching games generically have unique variable frame solutions (Janssen, 1997).

EXAMPLE 1, continued. To illustrate, consider Example 1 in detail. $\mathscr{C} = \{\mathsf{S}\}$, where S denotes the shape classifier. Hence, $\mathscr{F} = \{\emptyset, \mathsf{S}\}$. If $\mathsf{f} = \emptyset$, the only option is b, so trivially $\phi(\emptyset) = b$. If $\mathsf{f} = \mathsf{S}$, beliefs about frames become potentially important. To determine whether ϕ is a vfe, we ask whether, if a player hypothesizes that her coplayer is using ϕ, she maximizes her expected utility by conforming to ϕ. Using the decision just determined for $\emptyset$, the relevant expected utilities are

$$\begin{aligned}\mathscr{E}U(h \mid \phi, \mathsf{S}) &= v(\mathsf{S} \mid \mathsf{S})\big[\hat{\phi}(h \mid \phi, \mathsf{S}) + \hat{\phi}(b \mid \phi, \mathsf{S})/K\big] \\ &\quad + \big[1 - v(\mathsf{S} \mid \mathsf{S})\big]/K, \\ \mathscr{E}U(c \mid \phi, \mathsf{S}) &= v(\mathsf{S} \mid \mathsf{S})\big[\hat{\phi}(c \mid \phi, \mathsf{S})/(K-1) + \hat{\phi}(b \mid \phi, \mathsf{S})/K\big] \\ &\quad + \big[1 - v(\mathsf{S} \mid \mathsf{S})\big]/K, \\ \mathscr{E}U(b \mid \phi, \mathsf{S}) &= 1/K,\end{aligned}$$

where $\hat{\phi}(o \mid \phi, \mathsf{S}) = 1$ if $\phi(S) = o$, and $\hat{\phi}(o \mid \phi, \mathsf{S}) = 0$ otherwise ($o \in \{h, c, b\}$).

There are three vfe's, but (PD) selects hh as the unique variable frame solution given $v(\mathsf{S} \mid \mathsf{S}) > 0$ (i.e., provided our rational VFT player does not believe with certainty that her coplayer cannot perceive the shape difference).[8] (SD) plays no role in Example 1, but it is crucial to the VFT prediction for Example 2 (Colored Bottles), for the reasons we discussed in the Introduction.

3. THE VARIABLE FRAME LEVEL-n THEORY

We build upon the frame apparatus of VFT. The basic use to which frames are put in VFLNT is in defining the options and beliefs a player works with in confronting her choice problem. Accordingly, a player's frame is the set of families of properties or classifiers that she is *consciously aware of and pays attention to* as she contemplates a problem. It is often natural to identify these as the classifiers she *notices*. In a matching game in which the objects are physical and differ in terms of the classifiers color C and shape S, a player might notice both, in which case she would have the CS (color/shape) frame, or only color, in which case she would have the C frame, or only shape, or neither.

However, it is not sufficient that a player have the perceptual/conceptual ability to become aware of some classifier; it is necessary that she pay conscious attention to the properties that define it when contemplating her

[8] If $v(\mathsf{S} \mid \mathsf{S}) = 0$, then all three vfe's are variable frame solutions.

choice problem. There is ample psychological evidence that humans subconsciously filter out much of the information available to them as a means of coping with information overload (see e.g. Broadbent (1982) and Wundt (1874).[9] So it is probable that while an individual could distinguish each claret-shaped bottle if pressed to do so by specific questions or tasks, in a simple choice task these very subtle atomizing characteristics would be ignored. The outcome would be as if the individual could not perceive them. Thus, while she might be competent to use both classifiers C and S, she might pay attention only to S and hence has frame S instead of CS.

Viewing the frame f of a boundedly rational level-n player in this way, it is natural to assume further that her conception of her options is circumscribed by $\mathrm{opt}_{\mathsf{f}}$ and (in contrast to a VFT player) does not include the options generated by the proper subframes of f.

It is clear that at least in most matching games a player may have the potential to distinguish all the objects, for example, by focussing her gaze on each in turn, or by asking herself what each reminds her of. Such discriminations are often of a private nature. We can represent this case formally by saying that player i's frame contains a further classifier, the idiosyncratic atomizing classifier I_i, whenever she is consciously aware of and pays attention to such differences.

In the next two subsections we describe the main features of level-0 players, then those of players of levels 1 and higher. In Subsection 3, we formally derive the distribution of choices over objects at each level. In the last subsection we extend the model to incorporate further types of players, equilibrium players and worldly players.

1. *Level*-0 *Players*

In LNT (Stahl and Wilson, 1995), a level-0 (L0) player has no model of her coplayers that induces a belief and an optimal choice against this belief. Instead, an L0 player randomly picks an option in innocence and without bias. In Stahl and Wilson (1995), an L0 player's behavior is described by a uniform distribution over A. In the language of the present model, Stahl and Wilson (1995) implicitly assume that the L0 player's frame induces the discrete partition of A.[10] A direct application of LNT to Example 1 would predict a uniform distribution of choices for L0 players, and this would

[9]Wundt wrote: "The basic phenomenon of all intellectual achievement is the so-called concentration of attention. It is understandable that in the appraisal of this phenomenon we attach importance first and therefore too exclusively to its positive side, to the grasping and clarification of certain presentations. But for the physiological appraisal it is clear that it is the negative side, the inhibition of the inflow of all other disturbing excitations . . . which is most important."

[10]This is indeed plausible, given the payoff differences among the strategies.

also imply a uniform distribution of choices for all higher level players. LNT does not predict the Schelling competence.

In our VFLNT we retain the notion that an L0 player has no model of her coplayers and that she randomizes uniformly over her options, but we define her options explicitly with respect to her frame. Her option set is opt_{f}, and it is over these frame-induced options that the L0 player randomizes.

In the rationalistic VFT described in Bacharach and Bernasconi (1997), when a player has a frame f consisting of several classifiers she entertains the possibility that her coplayers might have a different frame, subject to (NR), which restricts the entertained frames of her coplayers to subsets of f. But this is a reasoning process about the coplayers and violates the notion of an L0 player. An L0 player has no ideas at all about her coplayers, and so none about her coplayers' frames.

The next step in developing VFLNT is to define, for each potential frame f, unique conditional probability measures, $\pi_0(o \mid \mathsf{f})$ and $p_0(a \mid \mathsf{f})$, which give the probabilities that an L0 player will choose option o in opt_{f} and object a in A . Invoking the principle of insufficient reason, we assume that she randomizes uniformly over her alternatives (her options). As in UFT, that conditionally on choosing an option she randomizes uniformly over the extension of that option. That is, for each $o \in \text{opt}_{\mathsf{f}}$ and $a \in A$,

$$\pi_0(o \mid \mathsf{f}) = 1/N_{\mathsf{f}}; \qquad p_0(a \mid \mathsf{f}) = \pi_0(o_{\mathsf{f}}(a) \mid \mathsf{f})/|E(o_{\mathsf{f}}(a))|, \tag{2}$$

where N_{f} denotes the number of options in opt_{f}[11] and $o_{\mathsf{f}}(a)$ denotes the unique option of opt_{f} that contains a in its extension.

Note that for any atomizing classifier I_i, if $I_i \in \mathsf{f}$ then P_{f} is the discrete partition, so (2) implies that $p_0(a \mid \mathsf{f}) = 1/|A|$ for all a. As L0 behavior is invariant to the idiosyncratic features of I_i, whenever an L0's frame contains such an atomizing classifier we will omit the subscript and denote it by I.

Since different individual players can have different perceptual/conceptual abilities and different subconscious information processes, they can have different frames. As in VFT, we take the probability measure on the set of frames $\mathcal{F}$ to be exogenous and the same for all L0 players; we denote it by $w(*)$. Then the population distribution of L0 players' choices over objects of A is given by

$$p_0(a) = \sum_{\mathsf{f} \in \mathcal{F}} p_0(a \mid \mathsf{f}) w(\mathsf{f}). \tag{3}$$

Even though L0 players are decision makers of very limited powers indeed, this model of their behavior yields some Schelling competence. Let us return to Example 1 and ask how the L0 players fare.

[11]Note that this number is equal to the cardinality of P_{f}.

EXAMPLE 1 (Bottles), continued. Consider again the simple Schelling game Bottles. We assume $\mathscr{C} = \{S, I\}$, where S is the shape and I is the idiosyncratic classifier. For notational convenience, label the hock bottle 1 and the others $2, \ldots, K$. We assume for simplicity that the frame is never empty, so that $\mathscr{F} = \{\mathsf{S}, \mathsf{I}, \mathsf{SI}\}$.[12] If the player notices the shape difference and ignores idiosyncratic differences, $\mathsf{f} = \mathsf{S}$ and $P_{\mathsf{S}} = \{\{1\}, \{2, 3, \ldots, K\}\}$. Therefore, $p_0(1 \mid \mathsf{S}) = 1/2$, and $p_0(a \mid \mathsf{S}) = 1/2(K-1)$ for $a \neq 1$. On the other hand, if the player does not notice the shape difference then $\mathsf{f} = \mathsf{I}$ and P_{f} is the discrete partition, so $p_0(a \mid \mathsf{I}) = 1/K$ for all $a \in A$. Finally, if the player pays attention to both shape and idiosyncratic difference, $\mathsf{f} = \mathsf{SI}$, and P_{f} is the discrete partition again, $p_0(a \mid \mathsf{SI}) = 1/K$ for all $a \in A$.

Then the aggregate distribution of L0 choices is

$$\begin{aligned} p_0(1) &= \big[w(\mathsf{S})/2\big] + \big[1 - w(\mathsf{S})\big]\big/K = (1/K) + w(\mathsf{S})(K-2)\big/2K, \\ p_0(a) &= \big[w(\mathsf{S})/2(K-1)\big] + \big[1 - w(\mathsf{S})\big]\big/K \\ &= (1/K) - w(\mathsf{S})(K-2)/2K(K-1) \qquad (a \neq 1). \end{aligned} \tag{4}$$

Thus, if $w(\mathsf{S}) \approx 1$, L0 players choose the hock bottle nearly half the time, and as long as $w(\mathsf{S}) > 0$ the hock bottle is the modal choice among L0 players. In this sense, the L0 players possess a limited Schelling competence. We shall see later that this tilt in the choice distribution of an L0 towards the hock bottle, though it may be slight, can induce a much greater tilt in the choices of higher level players.

EXAMPLE 2 (Colored Bottles), continued. The L0 behavioral predictions for this discrete Schelling game are exactly the same as for Example 1. To see this, note that the color classifier induces the discrete partition, and hence the same behavior as the idiosyncratic classifier I. Thus, the only probability parameter that matters behaviorally is $w(\mathsf{S})$, the probability that the L0 player uses the classifier S exclusively. Of course, it is not unreasonable to expect the *value* of $w(\mathsf{S})$ in Colored Bottles to be lower than in the previous example, because using the S frame entails being unaware of both the idiosyncratic differences and the color differences. Nonetheless, as long as $w(\mathsf{S}) > 0$, the hock bottle is more likely to be chosen by L0 players than any claret bottle, a prediction that VFT makes only after imposing both Payoff Dominance and Symmetry Disqualification.

[12] Ignoring the empty frame does not rule out any potential behaviors, since it would merely lead to uniformly random choices for L0 and L1 players, and this behavior is already possible as I is in $\mathscr{C}$.

2. *Features of Positive Level Players*

A level-1 player, unlike L0s, is a rational Bayesian decision maker, and as such she has a model of her coplayer that induces a belief from which she determines an optimal choice. In particular, she believes him to be L0 (as in LNT). In our VFLNT, we further assume that an L1 player has a frame f and believes that her L0 coplayer has a frame that is a subframe of f, as required by (NR). Since an L0's frame may in fact contain classifiers not in the L1's frame, the (NR) principle implies that the L1 may be mistaken. Equally, she may be mistaken in believing that her coplayer is an L0. She suffers from both "frame myopia" and "level myopia."

We give precision to our assumptions about an L1's model of her coplayer as follows. We begin by supposing that at some (sub)conscious level, perhaps very briefly, the L1 player reacts psychologically like an L0 player, but then conscious reasoning takes over. During this "L0 moment," some frame $\mathsf{f} \in \mathscr{F}$ is generated in the L1 player's mind. Just as for an L0, the L1's frame f is governed by an exogenous distribution, $w_1(*)$, say. In general $w_1(*)$ may differ from $w(*)$. For the moment, however, we assume for simplicity that $w_1(*) = w(*)$.

Her own frame, thus determined, is a major input into the L1's perception of options and of her beliefs. But now we need to specify the subjective probabilities of an L1 with frame f for her coplayer's frame; these we will denote by $\widehat{w}(* \mid \mathsf{f})$. We make two assumptions.

The first is just the (NR) principle: we assume that the support of $\widehat{w}(* \mid \mathsf{f})$ is a subset of $F(\mathsf{f})$.

Second, an L1's belief about *which* frame in $F(\mathsf{f})$ her coplayer has may be biased towards her own frame, f. If the L1 player has rational expectations (subject to (NR)), then $\widehat{w}(* \mid \mathsf{f})$ will be equal to $w(*)$ truncated to $F(\mathsf{f})$. On the other hand, it is reasonable to expect that an L1's model of an L0 player will suffer from "noticer bias": the tendency by an agent who has thought of a certain way of classifying objects to overestimate or underestimate the probability that others will classify them in the same way. Bacharach and Bernasconi (1997) found evidence of positive noticer bias. It is, for example, quite possible that the L1 player will simply project her own state of mind onto the other player; i.e., she will believe that the other player has frame f also, i.e., $\widehat{w}(\mathsf{f} \mid \mathsf{f}) = 1$. But this is an extreme case. A simple form of positive noticer bias of variable degree can be captured by one parameter in the following manner. Let

$$\begin{aligned} \widehat{w}(\mathsf{f}' \mid \mathsf{f}) &= \frac{b(\mathsf{f}' \mid \mathsf{f})\, w(\mathsf{f}')}{\sum_{\mathsf{f}'' \in F(\mathsf{f})} b(\mathsf{f}'' \mid \mathsf{f})\, w(\mathsf{f}'')} \qquad \text{for } \mathsf{f}' \in F(\mathsf{f}), \\ &= 0 \qquad \text{otherwise,} \end{aligned} \tag{5}$$

where $b(\mathsf{f}' \mid \mathsf{f}) = b$ if $\mathsf{f}' = \mathsf{f}$ and $b(\mathsf{f}' \mid \mathsf{f}) = 1$ otherwise, and $b > 1$ for positive noticer bias and $0 < b < 1$ for negative noticer bias. Formula (5) truncates $w(*)$ to the set of frames compatible with f and distorts it in the direction of any noticer bias.

So much for the beliefs of level-1 players. We turn now to their optimization problem. As stated before, the options of an L1 player with frame f are opt_f, the options given by frame f. For each such option, the L1 player has the necessary information to compute the expected utility from choosing that option. Stahl and Wilson (1995) provided evidence that players not only deliberate at different levels, but also deviate stochastically from the maximizing choices defined for the different levels above 0. These deviations may be interpreted as the result of computational error or idiosyncratic features. LNT assumes that deviations are less probable the greater the associated payoff loss is, according to a logistic choice function. Here we make the analogous assumption, Fallible Optimization (FO). The precision of the logistic function measures the degree of the player's fallibility. (FO) applies to all positive levels and we allow the precision parameter to vary with level.

Level-2 (L2) types are very much like L1 types except that they think their coplayer is either L0 or L1. Our L2 model is motivated by the following story of iterative thinking. An L2 player begins with an L0 moment during which a frame f is generated. Next, the L2 player starts thinking like an L1 player, forming beliefs $\widehat{w}(* \mid \mathsf{f})$ about L0 frames, which lead to a tentative probability measure for her coplayer's choice. As our L2 player computes the expected utility to this tentative measure, she realizes that some of the other players could do the same. She forms a belief about the proportions of L0 and L1 types in the population, α_0 and $1 - \alpha_0$. We represent these beliefs by a point estimate, $\widehat{\alpha}_0(2)$. As part of this next step, the L2 player realizes that some of these L1 players might not have the same frame as her, but perhaps some subframe in $F(\mathsf{f})$.

We assume that L2s have the same frame-conditioned model of L0s as L1s do. Like L1s, L2s may suffer from noticer bias in assessing their coplayer's likely frame, and this bias carries over into their model of how L1s think about an L0's frame (that is, they ascribe to L1s their own point estimates of frame probabilities). Given their frame, they get everything right except for, possibly, three things: this bias, the type distribution, and the L1 precision.

The behavior of L1s depends, as we shall see in the next subsection, upon their frame-conditioned probabilities $\widehat{w}(\mathsf{f}' \mid \mathsf{f})$ for their coplayers' frames, and that of L2s depends both on these probabilities and on their estimates of α_0 and L1 choice behavior. That of level-n players for $n > 2$ depends on analogous bundles of parameters.

3. *Positive-Level Play*

In this subsection we define the behavior and beliefs of level-n players ($n = 1, 2, \dots$) by a group of recursive relations, and show inductively why the behavior and beliefs so defined follow from the general features of positive level players that we have just described. We first recall that level-0 players' choices are given in (2) as

$$\pi_0(o \mid \mathsf{f}) = 1/N_{\mathsf{f}}; \qquad p_0(a \mid \mathsf{f}) = \pi_0(o_{\mathsf{f}}(a) \mid \mathsf{f})/|E(o_{\mathsf{f}}(a))|.$$

The statements (6)–(9) that follow describe an Ln player ($n = 1, 2, \dots$) with frame f. Let $\psi_n(o' \mid \mathsf{f}')$ denote the Ln player's belief that her coplayer will choose option o' when the coplayer has frame f'; for each f' in $\mathscr{F}$, $\psi(* \mid \mathsf{f}')$ is defined on $\text{opt}_{\mathsf{f}'}$ and this is its support. Then, the expected utility of any option o in opt_{f} given frame f and belief ψ_n is

$$\mathscr{E}U_n(o \mid \mathsf{f}) = \sum_{\mathsf{f}' \in F(\mathsf{f})} \sum_{o' \in \text{opt}_{\mathsf{f}'}} U(o, o') \psi_n(o' \mid \mathsf{f}') \, \widehat{w}(\mathsf{f}' \mid \mathsf{f}), \tag{6}$$

where $U(o, o')$ is defined by (1).

The Ln player chooses an option o in opt_{f} based on $\mathscr{E}U_n(o \mid \mathsf{f})$. We assume that the L$n$ player's choice follows $\pi_n(*)$, a probability distribution over opt_{f}. If $\mathscr{E}U_n(* \mid \mathsf{f})$ has a unique maximizing option in opt_{f}, say o_{f}^*, we might assume that the probability that the player chooses it is unity: $\pi_n(o_{\mathsf{f}}^* \mid \mathsf{f}) = 1$. But for greater empirical generality and in accordance with (FO) we shall allow $\pi_n(* \mid \mathsf{f})$ to be a logistic choice function of $\mathscr{E}U_n(* \mid \mathsf{f})$ with precision parameter γ_n,

$$\pi_n(o \mid \mathsf{f}) = \frac{\exp[\gamma_n \mathscr{E}U_n(o \mid \mathsf{f})]}{\sum_{o' \in \text{opt}_{\mathsf{f}}} \exp\left[\gamma_n \mathscr{E}U_n(o' \mid \mathsf{f})\right]}. \tag{7}$$

We specify an Ln player's ψ beliefs recursively, starting with $\psi_1(* \mid \mathsf{f}') = \pi_0(* \mid \mathsf{f}')$. For $n \geq 2$, let $\widehat{\alpha}_k(n)$ denote the Ln player's subjective probability that her coplayer will be an Lk player, for $k = 0, 1, \dots, n-1$. As in LNT, we assume that $\widehat{\alpha}_k(n) = 0$ for all $k \geq n$. For notational convenience, let $\underline{\widehat{\alpha}}(n)$ denote the vector of n parameters $\{\widehat{\alpha}_k(n)\}$ ($k = 0, 1, \dots, n-1$). Then ψ_n is given by

$$\psi_n(o' \mid \mathsf{f}', \underline{\widehat{\alpha}}) = \sum_{k=0}^{n-1} \widehat{\alpha}_k(n) \, \pi_k(o' \mid \mathsf{f}'), \tag{8}$$

where $\pi_k(o' \mid \mathsf{f}')$ is given by (7).

It is straightforward to derive the probability that an Ln player will choose any $a \in A$. As in (2) and (3),

$$p_n(a \mid \mathsf{f}) = \pi_n(o_{\mathsf{f}}(a) \mid \mathsf{f}) |E(o_{\mathsf{f}}(a))|^{-1}; \qquad p_n(a) = \sum_{\mathsf{f} \in \mathscr{F}} p_n(a \mid \mathsf{f}) \, w(\mathsf{f}). \tag{9}$$

The subjective probabilities $\underline{\widehat{\alpha}}(n)$ in (8) are implicit in $\mathscr{E}U_n$ and π_n. Hence, an L3 player must have not only subjective probabilities $\underline{\widehat{\alpha}}(3)$ but also beliefs about $\underline{\widehat{\alpha}}(2)$. For simplicity, we assume that, for all $n > 2$, the belief of an Ln player about $\underline{\widehat{\alpha}}(n')$ for $2 \leq n' < n$ is simply $\underline{\widehat{\alpha}}(n)$ truncated at n'; in other words, Ln players believe that a lower-level n' player's subjective probabilities are equal to the Ln's subjective probabilities conditional on $k < n'$. Further, when the choice probabilities π_n are specified as logistic with precision parameter γ_n, then beliefs about these parameters are also implicit in $\mathscr{E}U_n$ and π_n. Formally, let $\widehat{\gamma}_k(n)$ denote that belief of an Ln player about the precision γ_k of an Lk player, for $k = 1, \ldots, n-1$. Again, for simplicity, we assume that for all $n > 2$, the belief of an Ln player about $\underline{\widehat{\gamma}}(n')$ for $1 \leq n' < n$ is simply $\underline{\widehat{\gamma}}(n)$ truncated at n'.

4. *Equilibrium Players and Worldly Players*

For the purpose of empirically testing VFLNT it is very useful to have a model that encompasses both the pure theory under consideration and alternative theories. Since VFLNT is essentially a mixture of probabilistic choice functions over types of players, it can easily be extended to include an alternative theory by introducing a type whose behavior corresponds to the alternative theory.

To introduce a type corresponding to VFT, we define a "type E" player to be one who conforms to the equilibrium theory VFT. When there is a unique variable frame solution ϕ a type E player will conform to it, choosing option $\phi(\mathsf{f})$ when in frame f. Hence, the distribution of object choices by type E players is given by

$$p_{\mathrm{E}}(a \mid \mathsf{f}) = \left|E(\phi(\mathsf{f}))\right|^{-1} \text{ if } a \in E(\phi(\mathsf{f})), \qquad 0 \text{ otherwise.}$$

When there are multiple variable frame solutions, an E player conforms to some variable frame solution, and we assume that the subpopulation of type E players uniformly distribute themselves over all variable frame solutions.[13]

Similarly to LNT, we define a "worldly" type as a player who thinks her coplayer is either a type E player or else an L0 or L1 type.[14] Let $\widehat{\alpha}_0(\mathrm{W})$ and $\widehat{\alpha}_1(\mathrm{W})$ denote a worldly type's subjective belief as to the proportion of the

[13]In lieu of this latter assumption, one could restrict attention to cases in which there is a unique variable frame solution.

[14]The worldly type could be more generally defined as a player with the belief that all other players are either type E or "empirically significant" level-n types. Stahl and Wilson (1995) found strong evidence for E types and level-0 and level-1 types, but very weak evidence for level-2 types, and so specified the worldly type accordingly. Our specification could be revised in the light of new evidence.

population of coplayers that are L0 and L1 types respectively; a proportion $1 - \widehat{\alpha}_0(\mathrm{W}) - \widehat{\alpha}_1(\mathrm{W})$ is believed to be type E.[15]

4. VFLNT AND THE SCHELLING COMPETENCE

In this section we turn back to the phenomenon whose explanation motivates this paper. We ask when, and how, VFLNT explains Schelling competence, the repeatedly corroborated ability of real decision makers to solve coordination problems by matching on focal points, that is, by collectively alighting on a *salient* coordination point.

1. *Level*-1 *Play in Examples* 1 *and* 2

Let us apply the general scheme of level-n beliefs and behavior we have just developed to describe level-1 and level-2 play in our leading examples. In this way we will gain some insight into the processes by which VFLNT explains the Schelling competence, which differ sharply from those by which VFT explains it. In Subsection 2 we will give an exact condition under which the VFLN model yields focusing in a wide class of Schelling-like games.

EXAMPLE 1, continued. In the simple Schelling game Bottles, an L1 player can have three possible frames: S, I, or SI. Thus if she has frame I, she believes her (L0) coplayer has frame I also. In this case the coplayer has K options i_k, each an idiosyncratic option having a single object as its extension. Since the L0 coplayer chooses each of her options with equal probability, $\pi_0(i_k \mid \mathsf{I}) = 1/K$ $(k = 1, \ldots, K)$. Hence, from (8), the L1 player has subjective probability $\psi_1(i_k \mid \mathsf{I}) = 1/K$, so, from (6), all choices are equally good. Hence the choice probabilities $\pi_1(i_k \mid \mathsf{I})$ and $p_1(a \mid \mathsf{I})$ are also uniform, over $\mathrm{opt}_{\mathsf{I}}$ and A respectively.

On the other hand, if an L1 player has frame S, then she believes her (L0) coplayer has frame S also. But, from (8) above, her subjective probabilities for her L0 coplayer's option choice in this case are $\psi_1(h \mid \mathsf{S}) = 1/2$ and $\psi_1(c \mid \mathsf{S}) = 1/2$. Since $|E(h)| = 1$, and $|E(c)| = K - 1$, the best response is obviously to choose the hock bottle. For any positive precision $(\gamma_1 > 0)$, the modal choice will be the hock bottle, and as this precision increases $p_1(1 \mid \mathsf{S})$ approaches 1.

[15] Since the meaning of "having frame f" is different for type E players and for Ln players, there is no *a priori* reason why a worldly type should believe that type E players and Ln players have the same distribution of frames. Similarly, let $\widehat{w}(\mathrm{W})$ denote the worldly type's belief about the distribution of Ln frames. While $\widehat{w}(\mathrm{W})$ has the same functional role as $\widehat{w}$ in pure VFLNT, it is not necessarily the same distribution: i.e., worldly types and Ln types could have different beliefs about the distribution of Ln frames. That they be the same is a hypothesis that can be empirically tested.

The more complex case is that an L1 player has frame SI, because then she will contemplate whether the L0 player has frame S or I or SI. If the L0 coplayer has either I or SI, he chooses object a with probability $1/K$ for all $a \in A$. If the coplayer has S, he chooses h and c each with probability 1/2. Now opt_{I} consists of K options, $h \wedge i_1, c \wedge i_2, \ldots, c \wedge i_K$ (where i_k denotes the $L1$ *player's* kth idiosyncratic option. Hence the L1 player's expected utilities for her options are

$$\mathscr{E}U_1(h \wedge i_1 \,|\, \mathsf{SI}) = (1/K) + \widehat{w}\,(\mathsf{S}\,|\,\mathsf{f})\,(K-2)\,/2K,$$
$$\mathscr{E}U_1\,(c \wedge i_k \,|\, \mathsf{SI}) = (1/K) - \widehat{w}\,(\mathsf{S}\,|\,\mathsf{f})\,(K-2)\,/2K\,(K-1)$$
$$(k = 2, \ldots, K).$$

Given $\widehat{w}(\mathsf{S}\,|\,\mathsf{f}) > 0$, choosing the hock bottle is still optimal and hence will be the most likely choice of L1 players.

EXAMPLE 2, continued. In Colored Bottles, the L1 player's subjective probabilities for her coplayer's options, given the L1 player's frame f, are still the same as in Bottles, but now f ranges over $\mathscr{F} = \{\mathsf{C}, \mathsf{S}, \mathsf{I}, \mathsf{CS}, \mathsf{CI}, \mathsf{SI}, \mathsf{CSI}\}$. However, this extended set of frames generates the same set of partitions, and hence option extensions, as $\mathscr{F}' = \{\mathsf{S}, \mathsf{I}', \mathsf{SI}'\}$, where I' stands for I and/or C, so one could restrict the range of f to $\mathscr{F}'$ without altering the set of extensions of the options available in one frame or another. This implies that the only difference between L1 behavior in Bottles and Colored Bottles stems from the difference, if any, between $\widehat{w}(\mathsf{S}\,|\,\mathsf{SI})$ and $\widehat{w}(\mathsf{S}\,|\,\mathsf{SI}')$.

Level-2 play in these games is briefly dealt with.

EXAMPLE 1: Bottles, continued. We have just seen that the overall distribution of L1 choices peaks on the hock bottle (it is the modal choice under frames S and SI and the distribution is uniform under I). Hence the presence of $\pi_1(*\,|\,\mathsf{f}')$ in $\psi_2(*\,|\,\mathsf{f}')$ tilts ψ_2 in the direction of the hock option. This tilt will magnify the expected utility difference between the hock option and the claret option, and the L2 player will therefore be even more likely to choose the hock bottle.

EXAMPLE 2: Colored Bottles, continued. For the same reasons, the L2 player will, in this example too, be even more likely to choose the hock bottle.

Thus, like rational VFT, VFLNT predicts, both in simple Schelling games like Bottle and in discrete Schelling games like Colored Bottles, a shared bias toward a particular object and hence a measure of coordination. Since the object is the "unique possessor of a conspicuous feature" (it is the only hock-shaped one, and $w(\mathsf{S}) > 0$), it is salient in the sense of Lewis (1969) and Schelling (1960), and the coordination is thus on a focal point.

Here the similarity between the explanations ends. The root of the VFLNT explanation is that when there are obvious differences of some kind among the objects in A, there is a tendency for *level-0* players to define their options by those differences. This tendency is unthinking. It acts as a "seed" which propagates upward through the hierarchy. The VFLNT explanation of the population tendency toward the focal point combines this nonrational tendency at level-0 with rational capitalization on it by higher-level players.

Thus VFLNT explains focal play in a population by adding to the asymmetric strategic rationality of LNT an asymmetry in classificatory rationality. Level-0 players define their options in terms of a classification of objects which they employ without reflection. Higher level players not only classify but reflect about how others may classify; they recognize not only that coplayers define their options in terms of some classification, but also that this may differ from their own. VFLNT does not address the controversial question "Does pure rationality indicate choosing the salient option when playing with a purely rational coplayer?" (Gauthier, 1975; Heal, 1978; Gilbert, 1989); instead, it takes as its starting point the empirical existence of a hierarchy of less and less boundedly rational players and asks how such a hierarchy behaves in games with salient options. It builds on the simple idea that if some people choose a salient object without thinking then it is rational to do so oneself. But it gives a precise formulation to this imprecise idea by incorporating the model of agents' descriptions provided by VFT.

2. *A Focal Point Theorem*

In this subsection we show, for an arbitrary plain matching game, that if it has a strong enough focal point at an object a, then a is the modal choice for players of all levels. Moreover, as the focal strength increases, the choice distributions converge on the focal object. By saying that a focal point (o, o) is "strong" we mean that there is a frame f in which o is the only singleton option, no other frame defines a different singleton option, and the salience $w(\mathsf{f})$ of f is sizeable.

THEOREM. (Focal Point Theorem). *Suppose* (i) $K \geq 3$, (ii) *there are a frame* $\mathsf{f}^* \in \mathrm{Opt}$ *and an object* $a^* \in A$ *such that* $\{a^*\}$ *is the only singleton in* P_{f^*}, (iii) *for no other frame* f *does* P_{f} *contain a singleton different from* $\{a^*\}$, (iv) $b = 1$, (v) $\gamma_n = \infty \, \forall n$, (vi) $w(\mathsf{f}^*) \geq \max\{(K+1)(K-4)/(K+4)(K-3), \frac{1}{3}\}$. *Then* $p_n(a^*) > p_n(a)$ *for all* $a \neq a^*$, $n \geq 0$, *and* $p_n(a^*) \to 1$ *as* $w(\mathsf{f}^*) \to 1$.

The proof is given in the Appendix.

5. ASSESSMENT

1. *Contrasting VFT and VFLNT*

There is not much room to discriminate between the two theories VFT and VFLNT by comparing their predictions for matching games, since VFLNT is deliberately designed to emulate VFT's successful prediction of observed behavior in those games (in particular, the Schelling competence), and we have seen that it goes a long way toward doing so.

However, VLFNT predicts high Schelling competence only for sufficiently high level (and W and E type) players whose frames "express" the focal point. As we have seen in simple and discrete Schelling games, even those whose frames express the focal point have only a limited tendency to favor it at level 0 if they also notice idiosyncratic differences, and the focal tilt is still relatively weak at level 1. By contrast, subjects of VFT with such a frame choose the focal point with probability 1. This difference between the choice distributions implied by the two theories affords a discriminating test.

Second, there are possible means of directly testing the differing assumptions of the two theories.

We have emphasized that VFT and VFLNT differ in the rationality assumptions they make about players. VFT players are assumed to be highly rational in several respects and to have common knowledge of being rational in the same way. The way they are rational includes being Bayesian and having rational beliefs about each other's frames (subject to (NR)). In contrast, no level-*n* players model their coplayers as being like themselves, and indeed level-0 players have no model at all of their coplayers. Level-0 players are not even expected-utility maximizers. In this subsection we note that this last feature of level-0 players implies that in some games, called *hi–lo* games, their *framing propensities* are likely to differ from those of positive-level players or VFT players. The knock-on effects of this difference on higher-level players then give different distributions of choices in the VFLN and VFT models, affording a third discriminating test.

We take up in turn each of these three ways of empirically comparing VFT and VFLN.

1.1. *Distributional tests with classical matching games*. The basic framing effects which appear in VFLNT are the probabilities of the frames $w(\mathsf{f})$ and the frame beliefs $\widehat{w}(\mathsf{f}' \mid \mathsf{f})$. VFLNT can be tested empirically against VFT by a two-stage procedure. First, the framing parameters are estimated from behavior in ancillary tasks.[16] Then, using these estimates, the remaining

[16]In a simple *Buridan task* (BT) (Bacharach, 1996) a subject is presented with an object set A and asked to choose any one to obtain a fixed reward. In the *prompted* version (PBT) she

parameters of VFLNT are estimated from behavior in a variety of matching games. In particular, one obtains estimates in this way of the population fractions α_0, $\alpha_1, \ldots$ of players of different types, as in Stahl and Wilson (1995). Since one type of player (the E type) obeys rationalistic VFT, these estimated type fractions provide a test of that model against the present alternative.

1.2. *Tests of individual assumptions*. Tests of individual axioms of VFT are reported in Bacharach and Bernasconi (1997). The weak performance to date of symmetry disqualification was one motive for developing the present model. This and other assumptions of VFT can also be tested by a general paradigm in which one introduces variant types one by one into VFLNT. For example, to test (SD) we would introduce another version of type E that conforms to VFT except for (SD). Similarly, to test the rational expectations assumption of VFT, we would introduce a version of type E in which $\hat{v}(* \mid \mathsf{f})$ contains potential noticer bias, analogously to $\widehat{w}(* \mid \mathsf{f})$.

1.3. *Tests using hi–lo games*. In a simple Schelling game such as Bottles, $\mathscr{C}$ contains two classifiers, an atomizing classifier I and a classifier which partitions the objects of A into an oddity and a remainder. A *hi–lo Schelling* game is a simple Schelling game except that the prize is lower for matching on the oddity than on an object in the remainder.

From the point of view of variable frame theory, payoff differences among the objects play two roles in players' decision making: they may or may not be noticed and so used to classify objects and define options and, if they are noticed, they may or may not be motivators, making an option preferable because it promises a higher payoff. The idea that payoff differences might not be noticed, or might not motivate, may at first seem implausible, but this may be only because the usual presentations of games in game theory make payoff variations highly salient and because of the recent dominance of the assumption of expected utility maximization. What is true, however, is that any player who seeks to maximize payoff is, *ipso facto*, aware of payoffs. If we think of payoff differences among objects as a classifier—an element of $\mathscr{C}$—then any such player has this classifier in

is also prompted to think about the characteristics of the objects. The subject's opportunities to randomize are kept to a minimum. In a *Buridan estimation task* (BET) the subject is shown A and asked to estimate the percentages who make various choices in the BT (or PBT) with object set A. We suggest that agents choose in a BT with set A much as L0 players do in a matching game with set A (recall that L0 players do not think strategically). We could therefore use the relative frequencies of choices in a prompted BT with object set A to estimate $w(*)$ from (3). We also suggest that in a BET with set A subjects would respond in a manner that would reveal potential noticer bias. Using BT and BET data on set A and (5), it should be possible to obtain estimates of $\widehat{w}(* \mid \mathsf{f})$.

her frame with probability 1. This applies to the rational players of VFT and also to all players of level 1 or above in VFLNT. It does not, however, apply to L0 players. This gives rise to a difference between the choice distributions predicted by VFT and VFLNT for large enough payoff spreads. The key is that if there is a possibility, however slight, that an L0 notices neither the nonpayoff classifier nor the payoff differences and so picks an object at random, then choosing the low-paying oddity is a bad idea for higher level players, if it pays low enough.

For any matching game, let Y be the classifier of objects of A by payoff size: that is, Y puts a and a' in different cells if and only if matching on a and matching on a' give different payoffs. Consider a hi–lo Schelling game in which $\mathscr{C} = \{I, S, Y\}$. The above discussion implies that in the VFT model of this game

$$v(\mathsf{f}) = 0 \qquad \text{if } Y \notin \mathsf{f}, \tag{10}$$

and in the VFLNT model of the game, writing $w_t(^*)$ for the distribution of a type t player's frame over $\mathscr{F}$,

$$w_t(\mathsf{f}) = 0 \qquad \text{if } Y \notin \mathsf{f}\ (t \neq \mathrm{L0}).$$

EXAMPLE 3: Bottles with Payoff Differences. As in Example 1, there are K bottles ($K \geq 3$), one hock and $K-1$ claret, but now a match on the hock bottle pays only $y < 1$, while a match on a claret bottle pays 1. The classifier space $\mathscr{C} = \{I, S, Y\}$ as above, so $\mathscr{F} = \{\mathsf{I}, \mathsf{S}, \mathsf{Y}, \mathsf{IS}, \mathsf{IY}, \mathsf{SY}, \mathsf{ISY}\}$. But, by (10), in VFT the set of possible frames is $\mathscr{F}' = \{\mathsf{Y}, \mathsf{IY}, \mathsf{SY}, \mathsf{ISY}\}$.

Choosing the hock bottle is the unique VFT solution if $y > 1/(K-1)$. To see this, first observe that, by (10), $v(\mathsf{I}) = 0$, and so also $\hat{v}(\mathsf{I} \mid \mathsf{f}) = 0$ for all frames f that contain I. Thus, for the two possible frames containing I, viz. IY and ISY, the supports of coplayer subframes are $\{\mathsf{Y}\}$ and $\{\mathsf{Y}, \mathsf{SY}\}$ respectively. It is now straightforward to verify that if $y > 1/(K-1)$, payoff dominance (PD) picks a vfe in which both players choose the hock bottle. (There are two extensionally equivalent equilibria, since when a VFT player has frame ISY her option set contains the options *hock* and *low-paying*, both of which pick out the same bottle.)

In VFLNT, by (10), the set of an L1 player's possible frames is $\mathscr{F}'$. The aggregate L0 behavior $p_0(*)$ will be given by (3) with $w(\mathsf{S})$ replaced by $w(\mathsf{S} \cup \mathsf{Y} \cup \mathsf{SY})$. Hence, we easily find that

$$\mathscr{E}U_1(h \mid \mathsf{f}) = \left[\frac{1}{K} + \frac{\widehat{w}(K-2)}{2K}\right] y$$

and

$$\mathscr{E}U_1(c \mid \mathsf{f}) = \frac{1}{K} - \frac{\widehat{w}(K-2)}{2K(K-1)},$$

where $\widehat{w}$ is the subjective probability of the L1 player, given that she has frame f, that her coplayer's frame is a subframe of SY.[17] When $\mathsf{f} \in \{\mathsf{Y}, \mathsf{SY}\}$, $\widehat{w} = 1$, so $\mathscr{E}U_1(h \mid \mathsf{f}) > \mathscr{E}U_1(c \mid \mathsf{f})$—i.e., choosing the hock bottle is best—if and only if $y > 1/(K-1)$.

However, when $\mathsf{f} \in \{\mathsf{IY}, \mathsf{ISY}\}$, $\widehat{w} < 1$ is quite possible and reasonable. In this case, choosing the hock bottle is best iff $y > W/(K-1)$, where

$$W = \left[2K - 2 - (K-2)\widehat{w}\right] / \left[(K-2)\widehat{w} + 2\right],$$

which exceeds 1 for all $\widehat{w} < 1$. Thus, for an L1 with frame $\mathsf{f} \in \{\mathsf{IY}, \mathsf{ISY}\}$, if $\widehat{w}(\mathsf{S} \cup \mathsf{Y} \cup \mathsf{SY} \mid \mathsf{f}) < 1$ any value of y close enough to $1/(K-1)$ makes picking a claret bottle best. For example, if $K = 6$ as in Fig. 1, if $\widehat{w} = 0.95$ then picking a claret bottle is best if $y < 0.21$. In this case, the L0 tilt toward the hock bottle is not sufficient to overcome the lower payoff. For all higher level-n players, the tilt of the L1 player away from the hock bottle will be magnified. In sum, for frames f in $\{\mathsf{IY}, \mathsf{ISY}\}$, nonunit $\widehat{w}(\mathsf{S} \cup \mathsf{Y} \cup \mathsf{SY} \mid \mathsf{f})$, and low values of y, VFLNT, in contrast to VFT, predicts predominantly claret bottle choices. This is true independent of the fraction of L0's: even a very small fraction of L0's, if L1's suspect they notice atomizing differences, may suffice to persuade all higher-level players to go for claret.

A closely related game is the above game without the nonpayoff classifier S, so that the only nonidiosyncratic way of distinguishing the objects is that one pays less (y) than all the rest (1). VFT predicts the low-paying option provided $y > 1/(K-1)$, but casual empirical evidence suggests otherwise. One explanation of choices in the (potentially) high-paying remainder is that players try hard to find a differentiating classifier which they feel to be nonidiosyncratic, and sometimes do; this mechanism is discussed in Bacharach and Bernasconi (1997). Another is the above-VFLNT model. It is easy to see that this game is behaviorally equivalent to a hi–lo Schelling game in which S belongs to players' frames with probability 0. Since the above result does not depend on this probability, VFLNT predicts that many will choose objects in the high-paying remainder. We conjecture that some players, perhaps out of bewilderment at what seems a baffling problem of strategy, give it up and just choose at random one of the options momentarily salient to them. These would be true L0's.

2. *Concluding Remarks*

In the VFLNT we have presented here we have taken strategic rationality and framing propensities or saliencies to be essentially orthogonal, with $w(\mathsf{f})$ assumed independent of a player's level. But it may be asked: Would

[17]By (NR), $\widehat{w} = \widehat{w}(\mathsf{Y} \mid \mathsf{f})$ or $\widehat{w}(\mathsf{S} \cup \mathsf{Y} \cup \mathsf{SY} \mid \mathsf{f})$ according to whether $\mathsf{f} = \mathsf{Y}$ or SY.

not rational Bayesians also be rational framers, and do not rational framers notice all relevant aspects of a situation? In one way this question is misconceived: we hold that framing is logically prior to rational deliberation, for deliberation can take place only *within* a framework of concepts. However, there is a route by which high rationality may be associated with rich framing: we would expect rational decision makers to engage in a deliberate search for classifiers that might help them to structure their problems to suggest solutions, and in the case of coordination problems there is indeed a reason to think that this heuristic may help. There is therefore room for extending the current theory to allow for a rational classifier search. Moreover, it is empirically possible that people who reason deeply (or optimize surely) also tend to be good at perceiving situations under many aspects: we conjecture that they do.

The variable frame approach has theoretical potential in analyzing games for which there is strong evidence that option descriptions matter, and these include many major categories other than pure coordination games, including bargaining games (Roth and Murninghan, 1982; Mehta *et al.*, 1992) and social dilemmas (Evans and Crumbaugh, 1966; Colman, 1995). To date, the variable frame literature generally has concentrated on coordination games, but a few excursions into other classes of games shows its power. Within VFT, Janssen (1996) and Casajus (2000) apply it to bargaining games. The question arises as to what VFLNT might have to offer outside coordination games. We show elsewhere (Bacharach and Stahl, 1997) that it can explain puzzling findings of Rubinstein *et al.* (1995) in the zero-sum game of Hide and Seek. As in the analysis of hi–lo Schelling games above, our explanation exploits a plausible difference between the framing propensities of L0s and other types.

We have presented a variable frame level-n model of behavior in games that combines the fully rational variable frame theory of Bacharach (1993) and the boundedly rational level-n theory of Stahl and Wilson (1995). We have tried to offer a theory responsive to empirical data, both fitting existing observations and open to testing by future observations; and we have suggested strategies for testing it, which we are pursuing. Within the scope of the present essentially theoretical paper we have demonstrated that the variable frame approach has important and significant effects which are independent of the rationality assumptions of VFT. For the most part VFLNT makes the same qualitative predictions as VFT in matching games. Most notably, VFLNT explains the Schelling competence. But instead of relying on extended rationality concepts, the Schelling competence in VFLNT arises from the unthinking tendency of level-0 types to choose a salient object, and the boundedly rational deliberations of higher level types then magnify this initial tilt. The resolution of coordination indeterminacies offered by rational VFT appears to be more a product of the variable frame apparatus than of the rationality assumptions of that theory.

APPENDIX: PROOF OF FOCAL POINT THEOREM

We prove the Focal Point Theorem of Section 4.2.

THEOREM (Focal Point Theorem). *Suppose* (i) $K \geq 3$, (ii) *there are a frame* $\mathsf{f}^* \in \text{Opt}$ *and an object* $a^* \in A$ *such that* $\{a^*\}$ *is the only singleton in* P_{f^*}, (iii) *for no other frame* f *does* P_{f} *contain a singleton different from* $\{a^*\}$, (iv) $b = 1$, (v) $\gamma_n = \infty \, \forall n$, (vi) $w(\mathsf{f}^*) \geq \max\{(K+1)(K-4)/(K+4)(K-3), \frac{1}{3}\}$. *Then* (A) $p_n(a^*) > p_n(a)$ *for all* $a \neq a^*$, $n \geq 0$, *and* (B) $p_n(a^*) \to 1$ *as* $w(\mathsf{f}^*) \to 1$.

Proof. 1. We first establish claim A for $n = 0$. Write $\nu^* = |P_{\mathsf{f}^*}|$ for the number of cells in P_{f^*}. Then $p_0(a^* \,|\, \mathsf{f}^*) = 1/\nu^*$. Consider the subpartition of $A \setminus \{a^*\}$ induced by f^*, denoted $\widehat{P}_{\mathsf{f}^*}$. The probability of choosing some $a \neq a^*$ cannot exceed $1/2\nu^*$, since all cells in $\widehat{P}_{\mathsf{f}^*}$ are at least doubletons. In particular, if $K = 3$ then $p_0(a \,|\, \mathsf{f}^*) = 1/2$; if $K = 4$ then $p_0(a \,|\, \mathsf{f}^*) = 1/6$. Hence, $p_0(a^* \,|\, \mathsf{f}^*) - p_0(a \,|\, \mathsf{f}^*) \geq 1/4$ $(K = 3)$, $1/3$ $(K = 4)$. For $K > 4$, $p_0(a \,|\, \mathsf{f}^*)$ is greatest if a is in a doubleton cell, so $\widehat{P}_{\mathsf{f}^*} = \{\{a^*\}, \{a, a'\}, \ldots\}$; hence

$$p_0(a^* \,|\, \mathsf{f}^*) - p_0(a \,|\, \mathsf{f}^*) \geq 1/2\nu^*. \tag{11}$$

Next take any $\mathsf{f} \neq \mathsf{f}^*$ with no singleton cells except perhaps $\{a*\}$. If $\{a*\} \in P_{\mathsf{f}}$, then the above inequalities hold for $p_0(a^* \,|\, \mathsf{f}) - p_0(a \,|\, \mathsf{f})$. Otherwise, the worst case (that in which this difference is minimal) would be a binary partition in which a is in a doubleton cell and a^* is in the complement. Hence $p_0(a^* \,|\, \mathsf{f}) - p_0(a \,|\, \mathsf{f}) \geq \frac{1}{2}[1/(K-2)] - \frac{1}{4} = (4-K)/[4(K-2)]$, given $K \geq 4$; when $K = 3$, the only frame f with $\{a^*\} \notin P_{\mathsf{f}}$ and no other singleton cells is the degenerate partition $\{A\}$, in which case $p_0(a \,|\, \mathsf{f}) = 1/K \; \forall a \in A$. Therefore,

$$\begin{aligned} p_0(a^* \,|\, \mathsf{f}) - p_0(a \,|\, \mathsf{f}) &\geq 0 \; (K = 3, 4), \\ &\geq \frac{4-K}{4(K-2)} \text{ for } K > 4. \end{aligned} \tag{12}$$

Since $p_0(a) = \sum_{\mathsf{f}} p_0(a \,|\, \mathsf{f}) w(\mathsf{f})$,

$$p_0(a^*) - p_0(a) \geq \left[\begin{array}{ll} w(\mathsf{f}^*)/4 & \text{if } K = 3 \\ w(\mathsf{f}^*)/3 & \text{if } K = 4 \\ \dfrac{1}{2\nu^*} w(\mathsf{f}^*) - \dfrac{K-4}{4(K-2)}[1 - w(\mathsf{f}^*)] & \text{if } K > 4, \end{array} \right.$$

by (11) and (12), where $\nu^* = K/2$ for K even and $(K+1)/2$ for K odd.

Taking K *even* and $K > 4$, $p_0(a^*) - p_0(a) \geq [(1/K) + (K-4)/4(K-2)]$ $w(\mathsf{f}^*) - (K-4)/[4(K-2)]$, and this expression >0 iff

$$w(\mathsf{f}^*) > \frac{K(K-4)}{K^2-8}. \tag{13}$$

Taking K *odd* and $K > 4$, $p_0(a^*) - p_0(a) \geq [1/(K+1) + (K-4)/4(K-2)]$ $w(\mathsf{f}^*) - (K-4)/4(K-2)$, and this expression >0 iff

$$w(\mathsf{f}^*) > \frac{(K+1)(K-4)}{(K+4)(K-3)}. \tag{14}$$

If (14) is satisfied, then so is (13). This proves claim A for $n = 0$.

2. Next we establish claim A for $n = 1$. We have, for the probability of an L1 player that the L0 player chooses object a in frame f,

$$q_1(a \mid \mathsf{f}) = \sum_{\mathsf{f}' \in F(\mathsf{f})} p_0(a \mid \mathsf{f}')\widehat{w}(\mathsf{f}' \mid \mathsf{f}).$$

From Part 1, if $\mathsf{f}^* \in F(\mathsf{f})$ and

$$\widehat{w}(\mathsf{f}^* \mid \mathsf{f}) > \min\left\{\frac{(K+1)(K-4)}{(K+4)(K-3)}, 0\right\},$$

then $q_1(a^* \mid \mathsf{f}) > q_1(a \mid \mathsf{f})$ for all $a \neq a^*$. Further, since $b = 1$, it follows by (14) that $q_1(a^* \mid \mathsf{f}) > q_1(a \mid \mathsf{f})$ for all $a \neq a^*$ (and so $p_1(a^* \mid \mathsf{f}) = 1$) for all f such that $\mathsf{f}^* \in F(\mathsf{f})$. For any other frame f, since there are no singleton cells, $p_1(a \mid \mathsf{f}) \leq 1/2$. Therefore

$$p_1(a^*) - p_1(a) \geq 1 \times w(\mathsf{f}^*) - \frac{1}{2}[1 - w(\mathsf{f}^*)],$$

which >0 iff $w(\mathsf{f}^*) > 1/3$.

Claim (B) can be extended for $n > 1$; hence $p_n(a^* \mid \mathsf{f}^*) = 1$ for all $n \geq 1$. Moreover, as $w(\mathsf{f}^*) \to 1$, $p_n(a^*) \to 1$ for all $n \geq 1$. Q.E.D.

REFERENCES

Bacharach, M. (1993). "Variable Universe Games," in Frontiers of Game Theory (K. Binmore, A. Kirman, and P. Tani, Eds.). Cambridge, MA: MIT Press.

Bacharach, M. (1996), "When There's Nothing to Choose: A Model of Choice in Buridan Problems and an Experimental Design," Economic Beliefs and Behaviour Discussion Paper, University of Kent.

Bacharach, M., and Bernasconi, M. (1997). "The Variable Frame Theory of Focal Points: An Experimental Study," *Games Econ. Behav.* **19**, 1–45.

Bacharach, M., and Stahl, D. (1997). "Variable-Frame Level-n Theory with Application to Hide and Seek." working paper, Institute of Economics and Statistics, University of Oxford.

Broadbent, D. E. (1982). “Task Combination and Selective Intake of Information,” *Acta Psychologica* **50**, 253–290.

Casajus, A. (2000). “Focal Points in Framed Strategic Forms,” *Games and Econ. Behavior*, this issue.

Colman, A. (1995). “Game Theory and Its Applications in the Social and Biological Sciences,” Oxford: Butterworth–Heinemann.

Evans, G., and Crumbaugh, C. (1966). “Effects of Prisoner’s Dilemma Format on Cooperative Behaviour,” *J. Personality Social Psych.* **3**, 486–488.

Gauthier, D. (1975). “Coordination,” *Dialogue* **14**, 195–221.

Gilbert, M. (1989). “Rationality and Salience,” *Philosoph. Studies* **57**, 61–77.

Heal, J. (1978). “Common Knowledge,” *Philosoph. Quart.* **28**, 6–31.

Janssen, M. (1997). “Rationalizing Focal Points,” working paper, Department of Economics, Erasmus University.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard Univ. Press.

Mehta, J., Starmer, C., and Sugden, R. (1992). “An Experimental Investigation of Focal Points in Coordination and Bargaining: Some Preliminary Results,” in *Economics and Language* (W. Henderson, T. Dudley-Evans, and R. Backhouse, Eds). London: Routledge.

Mehta, J., Starmer, C., and Sugden, R. (1994). “The Nature of Salience: An Experimental Investigation of Pure Coordination Games,” *Amer. Econ. Rev.* **84**, 658–673.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice–Hall.

Roth, A. E., and Murnighan, J. K. (1982). “The Role of Information in Bargaining: An Experimental Study,” *Econometrica* **50**, 1123–1142.

Rubinstein, A., Tversky, A., and Heller, D. (1995). “Naive Strategies in Competitive Games,” working paper, The Foerder Institute for Economic Research, Tel Aviv University, 17 July 1995.

Schelling, T. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Stahl, D., and Wilson, P. (1995). “On Players’ Models of Other Players: Theory and Experimental Evidence,” *J. Econ. Behavior Organization* **25**, 309–327.

Sugden, R. (1995). “A Theory of Focal Points,” *Econ. J.* **105**, 533–550.

Van Huyck, J., Battalio, R., and Beil, R. (1990). “Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure,” *Amer. Econ. Rev.* **80**, 234–248.

Wundt, W. (1874). *Grundzuege der Physiologischen Psychologie* [*Principles of Physiological Psychology*]. New York/Berlin: Springer.