ISSN 1756-7904

DEV INTERNATIONAL DEVELOPMENT University of East Anglia

Working Paper Series

External validity in experimental methods: a social reality check

Cecile Jackson

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

UKaid from the Department for International Development

2011

UEA University of East Anglia

Working Paper 28

The School of International Development, University of East Anglia
Norwich, NR4 7TJ, United Kingdom

DEV Working Paper 28

External validity in experimental methods: a social reality check

Cecile Jackson

First published by the School of International Development in January 2011.

This publication may be cited as:
Jackson, C., 2011, External validity in experimental methods: a social reality check, Working Paper 28, DEV Working Paper Series, The School of International Development, University of East Anglia, UK.

About the Author
Cecile Jackson is a Professor of Development Studies at the School of International Development at the University of East Anglia, Norwich, UK.

Contact:
Email: Cecile.Jackson@uea.ac.uk
School of International Development
University of East Anglia
Norwich, NR4 7TJ
United Kingdom
Tel: +44(0)1603 593379
Fax: +44(0)1603 451999

ISSN 1756-7904

About the DEV Working Paper Series

The Working Paper Series profiles research and policy work conducted by the School of International Development and International Development UEA (see below). Launched in 2007, it provides an opportunity for staff, associated researchers and fellows to disseminate original research and policy advice on a wide range of subjects. All papers are peer reviewed within the School.

About the School of International Development

The School of International Development (DEV) applies economic, social and natural science disciplines to the study of international development, with special emphasis on social and environmental change and poverty alleviation. DEV has a strong commitment to an interdisciplinary research and teaching approach to Development Studies and the study of poverty.

International Development UEA (formerly Overseas Development Group)

Founded in 1967, International Development UEA is a charitable company wholly owned by the University of East Anglia, which handles the consultancy, research, and training undertaken by the faculty members in DEV and approximately 200 external consultants. Since its foundation it has provided training for professionals from more than 70 countries and completed over 1,000 consultancy and research assignments. International Development UEA provides DEV staff with opportunities to participate in on-going development work, practical and policy related engagement which add a unique and valuable element to the School's teaching programmes.

For further information on DEV and the International Development UEA, please contact:

School of International Development
University of East Anglia, Norwich NR4 7TJ, United Kingdom
Tel: +44 (0)1603 592329
Fax: +44 (0)1603 451999
Email: dev.general@uea.ac.uk
Webpage: www.uea.ac.uk/dev

Abstract

The relevance of experimental games as methods in development research depends crucially on how far the results from the games can be extrapolated to real life, ie the external validity of those results. The extent to which external validity matters depends on what you want to do with the data, some kinds of theory testing can arguably afford indifference, but many experiments are used as an indicator of behaviour in everyday life. This paper is focused on the 15 cross-cultural studies of Henrich *et al* (2004) as the most systematic attempt by anthropologists to use experimental games in developing countries, to assess the extent to which such methods can illuminate norms and social preferences in reality.

Key words: experimental methods, anthropology, gender, identities, developing countries, research methodology.

Acknowledgements

This paper is written for "The Intra-Household Allocation of Resources: Cross-Cultural Tests, Methodological Innovations and Policy Implications", a project jointly funded by the UK's ESRC and DFID (RES-167-25-0251).

Engagement by anthropologists with the experimental methods of economics suggests a promising disciplinary partnership; the extension into cross cultural studies and a deepening of contextual understanding for economics, and into new forms of observation-based methods for anthropology. The set of 15 studies under MacArthur funding (Henrich *et al* 2004)[1] represent the best known and most influential of these, and they come to at least two surprising conclusions which stimulate the question of how experimental results relate to the social reality external to the experiment.

The component studies are on the hunter gatherer and horticultural Conambo of Ecuador (Patton 2004), hunter gatherer and horticultural Machiguenga (Peru), subsistence farming Mapuche and wage working Huinca (Chile) (Henrich and Smith (2004), hunter-gatherer Hadza of Tanzania (Marlowe 2004), forager-horticultural Tsimane' of Bolivia (Gurven), foraging and horticultural villagers in Papua New Guinea (Tracer), pastoral-nomadic Mongols and Kazakhs in Western Mongolia (Gil-White 2004), Shona farmers in Zimbabwe (Barr, 2004), agro-pastoralist and farming Sangu of Tanzania (McElreath 2004), pastoralist Orma of Kenya (Ensminger 2004), hunter gathering Ache of Paraguay (Hill and Gurven 2004) and Lamalera whalers of Indonesia (Alvard 2004). These studies are placed on an evolutionary scale from hunter-gatherers to complex agriculturalists, and are focused on variations in cultures of altruism and selfishness. The anthropological authors are derive from evolutionary anthropology and work largely with an explicit or implicit evolutionary psychology hypothesis that success has accrued to groups able to develop norms as guidelines for sharing, which fostered advantageous cooperation. The Henrich *et al* analysis compares the results from Dictator, Ultimatum and Public Goods games[2] with those in the west, and the first surprising conclusion is that overall, higher levels of market integration are positively associated with 'fairer', less selfish behaviours. They argue that impersonal, large-scale, market-integrated societies require strong sharing and fairness norms because second-party enforcement is not possible, and this is why western subjects display relatively unselfish sharing in

[1] These studies were all part of a programme which conducted the same games – Ultimatum, Dictator, and Public Goods Games – in 15 societies arranged in a sequence from simplest hunter-gatherers to complex agriculture.

[2] These are social preference games, played in anonymous pairs with real money. In the Dictator game the proposer (who has the money) simply decides what amount to offer to the responder, and if completely self-interested should offer nothing. In the Ultimatum Game, the responder also decides whether to accept the offer, and if the offer is rejected then neither player receives any money. Self-interested players should make low offers, and responders, few rejections. Public Goods games involve players with money endowments making decisions whether to contribute to a group project which benefits all, even the non-contributors, and self-interested players should contribute nothing. (See Camerer and Fehr 2004: 60-78, for further details on games).

these games. The second surprising conclusion is that individual characteristics like gender have no, or very limited, influence on experimental behavior.

Given the accounts of simple societies as eqalitarian and community oriented, compared to individualistic, selfish, competitive western cultures, it was unexpected to find that, although Ultimatum games showed considerable variation[3], 'The self-interest axiom accurately predicts responder behaviour for about half of our societies, even though it generally fails to predict the responder behaviour of university students.' (Henrich 2004: 24). Many of the societies in the MacArthur studies are known for the high degree of sharing, compared to western societies, yet their members play the games more selfishly than western subjects. In explaining why societies with a high degree of sharing nevertheless exhibit selfish game play, Ensminger suggests that 'self-enforcement need not develop when second party enforcement is ever-present' (2004:359). This seems somewhat improbable, since second party enforcement is an ineffective approach, involving massive transaction costs, however small and face to face the group, as anyone who has ever had to enforce domestic labour by teenagers will know. Further, the implication that small-scale societies have fewer norms feels equally improbable.

The second surprising conclusion of Henrich *et al,* is that individual characteristics (occupation, age, gender etc) do not explain game behaviour variations. Whether markets undermine the moral foundations of society, as Marx thought, or alternatively, stimulate honesty and fair dealing, we should surely see distinctive behavior in the games from those individuals most involved in markets. However, we do not, overall, see this connection at individual level[4], and exactly how the evolution of unselfish norms in market relations is supposed to evolve is unclear.

One would expect that individuals with very different social positions and life experiences, women, men, the young or elderly, the rich, the poor, educated and uneducated, those with different livelihood bases, and so on, would relate to the sets of norms supposedly at play in relation to particular games in distinctive ways, and this would be reflected in how they played the games. The factors considered in the MacArthur studies are sex, age, education, and wealth, and in pooled regressions on all offers none of these attributes predicted offers[5].

[3] Variability is higher in these studies than among western subjects – and one PNG society rejected both very low and very high offers with equal frequency (Henrich *et al* 2004: 19).

[4] Although Ensminger for the Orma finds that those engaged in wage labour and trade were more likely to engage in fair behaviour than subsistence producers, which she sees as related to the importance of reputation in market oriented societies.

[5] But this 'does not exclude the possibility that the individual differences we have measured may predict behaviours in different ways from group to group' (Henrich 2004: 28).

First, how exactly was the connection between game behavior and external reality conceived? Henrich *et al* are 'convinced that local economic and social structures are reflected in the experimental behaviors we observed' (2004: 48) and offer two possible ways in which this occurs. One is that environments produce distinctive and generalized behavioural dispositions which apply across many domains, so that one group may be generally more altruistic than another. This however would imply similar behavior in differing games which is not evident in the component studies. And that similar environments would produce similar dispositions which is also not evident. Their second, and favoured pathway, is to see games as cueing highly context-specific behavioural rules, in which 'our subjects were first identifying the kind of situation they were in, seeking analogs in their daily life, and then acting in an appropriate manner.' (2004: 48). Whilst recognizing that it is difficult to distinguish which of these two pathways explain the results, the case of the Orma in Kenya is used to argue in favour of the context-specific trigger of norms governing everyday life. 'The Orma' are said to have made a connection between the Public Goods game and *harambee*, a local cooperative institution which wealthier members contribute more than poorer ones, and in the game behavior also found higher contributions from wealthier members. In the account of study methods for the Orma research, Ensminger makes it clear, however, that it was not the Orma subjects but the research assistants who made the connection between the public goods game and *harambee* ('when the game was first described to my research assistants they immediately identified it as the *harambee* game' (2004:376). If a game is labelled in this manner, the framing is so complete that it would be surprising if players did not do what was expected in *harambee*. Neither the 'generalised disposition' or the 'context-specific norms' pathways, suggested as the basis of external validity, stack up against the many contradictions and questions evident in the component studies, and discussed further below.

These issues prompted a questioning of how researchers using experimental methods connect game behaviour evidence with the social reality they are analyzing, ie the issue of external validity. How is the presence and strength of sharing norms, measured in the Dictator and Ultimatum games, extrapolated to external reality? And how are social identities, known to be important in patterning behavior, left at the door of the lab, and unimportant in game decisions? I will first discuss the absence of connection of experimental results to individual social identity, with a focus on gender, before considering what approaches would help understand the variable degrees of external validity which experimental games might hold. For this purpose, the framework of Levitt and List (2007) is used to set out the factors which are seen by economists as distorting the external validity of game results, and which is then extended with a discussion of the deeper contextual and theoretical issues which a social anthropological sensibility might see as important for the claims made form experimental game results.

1. Gender identities and game behaviour:

The big differences found in the MacArthur game behaviours were between, not within, groups, and the overall conclusions about market integration and sharing norms are arrived at through ranking of societies in an index composed of 'payoffs to cooperation' ie the extent of reliance on cooperation outside family level; 'market integration' ie extent of dependence on markets; 'anonymity' ie the extent to which people interact with strangers they may never see again; 'privacy' ie how possible it is to keep secrets from others; and 'sociopolitical complexity' ie how much centralised decision-making occurs above the level of the household.

Evolutionist explanations require some kind of material advantage to accrue to particular social behaviours (eg cooperation) in order for them to be selectively favoured through reproductive success[6]. But how exactly do some 'lifeways' prove more adaptive and successful, if not through the changed behaviours of individuals, for example traders and wage labourers within any particular group displaying norms and preferences in the games appropriate to market integrated societies, ie less selfishness and greater willingness to punish? Or, if exposure to markets and market relations leads people to make higher offers and more frequent rejections, one would imagine that women in simple societies, who are usually much less market integrated than men, would make lower offers and fewer rejections. But it appears that they do not. We see little or no evidence of changed behaviours of individuals within groups.

It is surprising that age, sex, education and relative wealth do not affect offers or the likelihood of a subject rejecting an offer. With a few group-specific exceptions, no individual characteristic other than group membership (village, camp etc) predicted experimental play. And even though market integration has such a major effect at group level this does not emerge in individual level measures – the behaviours of players with more or less involvement in wage labouring or cash cropping is not distinctive.[7]

Gender identity is generally treated rather casually by the researchers in the MacArthur studies, eg Henrich and Smith (2004) remark in an aside that their generalisation about the absence of livelihood cooperation for the Mapuche is not

[6] Eg Gil-White says that the puzzle of players overestimating the rejection rate is explained by their neuroses about their reputations and the fear of being seen as a bad person, despite the fact that nothing much seems to befall such bad people. This in turn is explained by ecological adaptation to low population density and herding livelihoods.

[7] In the one case where it did, the effect was that offers were higher amongst wage labourers) (Henrich *et al* 2004:35)[7].

true for female headed households, but there is no pursuit of this, suggesting a rather selective account of the cooperation context. In addition, there is not a discussion of what is labelled as 'cooperation', what is identified as sharing, and so on. Marriage is an important and variable form of cooperation in all societies, with very different kinds and levels of gender and generation cooperation in production and consumption, and in self-interest and altruism, but domestic intrahousehold cooperation is excluded from accounts of cooperation contexts and cultures. There are deeper difficulties too; the identification of some behavior as 'pro-social', or 'other-regarding', is part of a dichotomized perspective on social worlds which is difficult to reconcile with more complex anthropological views of societies as held together through conflict, or of self-interest as including the wellbeing of others.

It also seems curious that gender identity does not emerge as a significant factor in these game results, given the extensive work showing how deeply gender influences research encounters, and that women show distinctive preferences in 'mountains of research in psychology and sociology' (Eckel 2008) where they appear to be more altruistic and cooperative and less competitive than men. In a comprehensive review of experimental game results and gender in western societies, a complex picture emerges, but a strong case is made for the greater sensitivity of women to framing effects and context (Croson and Gneezy 2009), and Eckel's review (2008: 10-12) shows that women give more to anonymous counterparts in Dictator games, are more likely to accept offers in Ultimatum games, and men are more likely to punish unfair offers, while women's offers are more likely to be accepted as fair by both women and men.

There were moments in particular studies where gender was statistically significant in some elements of the games – eg in minimal acceptable offers to Conambo (Ecuador) women, where women accepted lower offers than men (Patton 2004:107)[8]. In the Bolivian study men offered, on average, rather more than women (Gurven 2004:215) and middle-aged women were significantly higher contributors to the public goods game. In the Ache study all the high Ultimatum Game offers were made by women, but whilst many of the highest Public Goods contributions are made by women the statistics finally showed men making significantly higher contributions than women (Hill and Gurven 2004: 398). In PNG Tracer (2004: 248) finds women's offers higher than men's, and women less likely to accept both very high and very low offers than men and, while not statistically significant, he concludes that this supports Eckel and Grossman's (1998) dictator game findings amongst western subjects that women tend to be more generous and less individually orientated than men. But in the Hadza study where gender is almost significant (Marlowe 2004: 185) in predicting offers, women as dictators offered less. In a number of studies gender appears at first glance to differentiate behaviour, but

[8] In this study, although sex was as important as location/identity, the analysis does not follow it up.

this generally all comes out in the statistical wash, and gender was rarely significant in explaining behaviour.

The absence of significance of individual characteristics like gender seems surprising because norms of appropriate behavior are so frequently gender differentiated, in relation to divisions of labour as well as symbolic constructions of gender. If a division of labour excludes women from, say, hunting game, then why would a norm for sharing hunting meat be exercised by women as well as men? One might argue that there are meta norms drawn on by women and men operating above the level of particular contexts like the hunt, but the analysis offered is in terms of very context specific norms, where one might be surprised that both genders equally deploy such a norm. The experimental anthropologist's fondness for explaining game play as deriving from norms which are grounded in the micro material realities of the cultures in which they are embedded feels at odds with insignificance of individual characteristics. For example, education of individuals in the MacArthur studies does not prove significant in explaining play, yet it most certainly alters norms and aspirations.

The absence of significance of gender in game play is also at odds with the external realities described. Most authors describe marked gender divisions of labour in the cultures studied, and Tracer, in the PNG study, describes the cultural environment as one in which gender marks strong divisions (men and women sleep separately, women are prohibited from spirit and meeting houses, men are fearful and anxious in relation to women who cause illness, sap energy, and diminish hunting ability, 2004: 238-9). One might therefore expect women and men to connect to dominant norms rather differently.

In addition to the connections between specific individual's livelihoods and game play, we might expect broader gender effects on play from society-wide gender ideologies and socialisation. Whilst puzzling over the contradictory results of his games in general, Gil-White remarks that 'the logic of the game may be easier to grasp for those who are socialised into the proposer role' (2004: 274). If women are connected to the prevailing male dominated cultural vocabularies and norms in subaltern ways then we would expect some reflection in game behaviour, if indeed game behaviour does have external validity. Certainly we might expect that women, who are generally less likely to be doing the proposing in relation to resource sharing, would find it harder to grasp the game and may make offers more divergent from the prevailing norms, yet Gil-White like others finds no significant gender effects on play.

After dismissing the significance of individual characteristics, the MacArthur authors look at group differences, which are felt to influence play more clearly than individual differences, yet these also produced some surprising results. Group effects

were found to be strong even where groups are not geographically isolated, thus the Quichua and the Achuar of Ecuador interact and intermarry but played the Ultimatum Game very differently (2004:36)[9]. And unexpectedly, given other experimental evidence on the readiness with which groupness affects behaviour in western subjects, Gil-White finds no in-group favouritism in offers and acceptances in the Mongolian study (2004: 281).

The results on group differences however leads Henrich to the view that 'Group-level measures of economic and social structure statistically explain much of the between-group variation in experimental play. This suggests that there might be a relationship between behaviour in our games and common patterns of interaction in daily life.' (2004:38). He goes on to say that '[B]etween group behavioural differences…. are the product of the pattern of social and economic interaction that frame the everyday lives of our subjects (2004:45)'. But why and how? If the everyday lives and patterns of social interaction of individuals do not explain anything, then how do group attributes? If it is some functional material reality which is at the root of all these explanations then how is it, by the same logic, that the specific material circumstances of individuals have no effect on play?

To summarise; the absence of significance of individual characteristics on play is puzzling and begs a number of questions. How are the connections between quotidian realities, norms and game play supposed to work if the functional logic fails? Has enough attention been paid to forms of conjugal and intrahousehold cooperation? There is a lot of evidence that women and men relate to hegemonic norms differently at a number of levels; livelihoods and material realities, relations to money and markets, subjective conceptions of rights and personal self-interest, behavioural ideals, and the symbolic vocabulary of their worlds and lifeways. Why does this find no reflection in play? These puzzles suggest a closer look at the claims for external validity of experimental game results.

2. External validity in experimental economics:

Amongst mixed discipline audiences, encountering experimental methods afresh, the first question always raised is what relevance the behaviour in experimental games has for everyday social behaviour, since the experiment is such an artificial situation. Oddly, to a social anthropologist, the relationship between experimental results and external reality has not excited very much interest in economics, perhaps because the ideal of the science laboratory is so methodologically pervasive, and in many sciences objects behave the same in a lab as outside it. Even where a gap between

[9] In this comparison the cooperative and market experienced Quichua played *less* generously than the Achuar.

laboratory and real life is recognised, Guala remarks that '[e]xperimental economists … tend to ignore or downplay the issue of external validity', and concentrate on experimental design (2005: 142) for several reasons. One argument is that where experiments are used to reject theories it does not matter how realistic they are since it is assumed that a failure of theory to predict in a lab situation will entail a certain failure in reality. This however assumes that the experiment is an adequate test *by itself*, to allow theory rejection, a confidence which would not be shared by many non-economists since the experiment may be at fault not the theory. And furthermore, the connection to reality of positive results from experiments is not helped with this position. Finally, the preoccupation with the dirty test tube problem, ie with ensuring that experimental results are not spoiled by design errors confounding analysis may also have worsened improve external validity, since there is a trade-off between internal and external validity in experimental design; 'The more artificial the environment, the better for internal validity; the less artificial, the better for external purposes.' (Guala 2005: 144).

A useful framework for considering how far insights gained from lab behaviour can be extrapolated to the real world has been proposed by Levitt and List (2007) who are interested, in their discussion of the class of experiments dealing with social preferences, in what 'distorts' lab behaviour, and who categorise distortions limiting external validity, into 5 areas, which are discussed next before raising further methodological questions about external validity from an anthropological stance.

2.1 Moral and ethical considerations

Decisions made by individuals reflect both the financial costs and benefits of alternative actions, and the moral costs or benefits of those actions. A decision at odds with a person's moral code or sense of identity carries costs, and the stronger the norm involved the more likely it is to dominate the financial pay-off, and conversely the greater the financial pay off the more likely it is for moral concerns to be set aside. Levitt and List conclude that 'games where the moral and the wealth-maximising choices are in conflict (eg experiments measuring social preferences) are likely to yield experimental results which diverge most sharply from real world behaviour.' (2007:29). These are the games that social development researchers are likely to be most interested in, and yet they are the most vulnerable to distortion effects. In addition, the balance of this trade-off is much less knowable for non-western cultures where different moral codes are at work, and where the value of any particular financial pay-off is also uncertain. Experiments on social preferences in developing countries therefore should carry a particularly stringent health warning.

2.2 The nature and extent of scrutiny

Lab behaviour is subject to scrutiny which exaggerates pro-social behaviours, eg hand washing in lavatories is more common when people think themselves observed. The same people that display what look like pro-social preferences (eg donations to charity) in experimental conditions, are found not to behave in similar ways outside of the experiment. However, the significance of this depends on the type of experiment, and scrutiny is seen as 'a minor problem in many experiments, especially if the decision environment is interactive and 'rich'' (Falk and Heckman 2009: 7).

Scrutiny may bias towards pro-social behaviour (Hoffman 1994), and even signals like eye-shaped images on instruction forms increases cooperative behaviour (Hayley and Fessler 2005). However removing the perception of scrutiny is very difficult, and even if this is possible, there is a great deal of variability of pro-social behaviour outside of experimental conditions, and little evidence of general consistency of behaviour – one context may stimulate a display of pro-social preferences and another may not – such that it is far from clear which norms are responsible for pro-social behaviour in an experiment. Even more challenging is the evidence for the temporality of pro-social acts in everyday life, discussed further below.

Anonymity (both of researcher and players, and between players) is a general rule in experimental methods. It affects behaviour, since less anonymity produces more pro-social behaviours. Anonymity of play is expected to remove the possibility that players behave in ways which they think they ought to, and mindful of their reputations. It has however been problematic in many of the MacArthur studies; in the Ache games in Paraguay (Hill and Gurven 2004: 394) the pay-offs were not really private and everyone knew that the (very well known) researcher knew how they had played, Alvard in the Indonesian study notes that talk between players was impossible to control, before and after the games, McElreath says that players often announced their offers or contributions on leaving the room (2004: 340), the Ache are unfamiliar with 'the possibility of anaonymous behaviours without social consequences (Hill and Gurven 2004: 402) and behaved similarly in public and private versions of the Publid Good games, and Ensminger for the Kenya study, points out that where guaranteed anonymity is very rare in everyday life, people may behave as if it does not exist.

The absence of a belief in the anonymity of the game may be part of why reputation management seems to be an important element in some game play, even though this does not necessarily mean behaving altruistically. Henrich and Smith's players all knew each other and the researcher well, so may have had reason to question anonymity, yet they still cooperated less than the US students interacting with

strangers (2004: 156). This prompts a questioning of the ethnocentrism of the experimenters belief that anonymity and privacy, based in western dichotomies of public and the private, and grounded in western lifeways and histories. Behaviour frontstage and backstage, in Goffman's terms, is very clearly delineated within western cultures, but possibly less, or differently, so, in others. Both Tanzanian Hadza (Marlowe 2004), and the PNG Au and Gnau (Tracer 2004: 240) were said to try to avoid sharing meat from hunting by sneaking into the village under darkness, and the former have so little everyday privacy that trousers are valued by Hadza men for having pockets which allow things to be kept in secret (Marlowe 2004:188). In these circumstances a game offering a treasured opportunity not to share cannot be seen as reflecting everyday norms. Marlowe argues that it is different kinds of privacy in complex societies which allow us to limit sharing; the Hadza share because of the impossibility of concealment.

Where very little is private, secret or anonymous, the stripping out of reputation-oriented decisions is not so easy. Another factor is aversion to lying. Privacy should allow people to act in ways which they may not wish to defend publically, by lying with impugnity. But lying well, with people you know intimately, is not easy: in discussion after the game Ache would have to lie convincingly to relatives and friends, and 'would be very uncomfortable at the thought of having to lie under these circumstances, and might therefore make fair offers because they knew they would be asked about their offer after the game ended.' (Hill and Gurven (2004: 404)

Even if player anonymity is secure there is nothing to prevent a player imagining their partner, as a consequence of the game framing, or simply creating a projection – as in the Zimbabwean woman player, who explained her acceptance of a zero offer by saying that only someone in dire need would have made such an offer, and so she had accepted it (Barr 2004: 319).

Economists and anthropologists may see the public and the private rather differently. Why should we expect that anonymity will make people behave completely differently in public and private, and might this have been exaggerated? The fact that 'no-one will know' does not cancel out the *social* elements of personal identities. People enact their cultures even when alone. What is social in us resides in our personhood, and is not only performed when others are present. Toilet users may be less likely to wash their hands when unobserved, but they are very likely to observe many other conventions, like closing the door. Furthermore, the self is part of the audience for behaviour. How we feel about ourselves is an ever-present check on how we act.

2.3. Framing and context

The classic work in economics on framing emphasises the ways that information may be coded positively or negatively by the words used to convey it (Tversky and Kahneman 1981), called valence framing, for example, in a risk experiment whether one talks about the risk of lives lost or lives saved induces different levels of risk aversion. Levin at al (1998) review literature on framing and suggest a typology of 'risky choice framing' referring to how the outcomes are described, 'attribute framing'[10] referring to the manipulation of a characteristic of an event or object in the game, and 'goal framing'[11] referring to the way the goal is labelled (1998: 150) in order to show systematic effects of types of framing. This is useful but the range and depth of framing recognised by social scientists goes well beyond positive and negative framing (Jackson 2009, 2010).

More recently, Benz and Meier comment on context more clearly by observing that 'Social norms may be triggered very differently in context-free and context-rich environments' (2008:269). The question of how much to cue the domain which one is researching - in the way the experiment is explained to players - is tricky; too little and you may not be triggering the norms you are interested in; too much and you may induce expected behaviour (Charness and Kuhn 2010). However these debates focus on how the experimenter knowingly manipulates the frame, and not on the unknown framing brought into the experiment by the players.

In their discussion of framing Levitt and List refer to the Henrich studies and seem to both accept the external validity claim yet also remark on the degree of uncontrolled framing by their subjects; if 'an experimenter mistakenly assumes that the agent is treating the game as one-shot then reputation building behaviour can be misconstrued [by the experimenter] as social preferences' (2007 :163). But if subjects are not playing the game the researchers think they are playing, then what meaning do you give to the results, let alone accepting that these results are consistent with everyday behaviour? Western researchers running experiments in other cultures, cannot really see their experiments on a continuum of 'context-free' to 'context-rich' as Benz and Meier suggest, but more realistically is actually context-ignorant, or context-confused, and raises major questions for the interpretation of results.

The creativity with which players frame games for themselves is little commented upon. Alvard however observes that ''Experimenters simplify the real world, create

[10] For example, meat is more likely to be evaluated positively if labelled '75% lean' compared to '25% fat'.

[11] For example, framing the same scenario outcome differently shows subjects are more willing to go accept breast self-examination if the same outcomes are described negatively compared to positively (1998: 168), ie negative goal framing is more persuasive.

the rules of a game, and have faith that their subjects play by them. Players may view the experiment in a different way – as a game within a game…the larger game does not end with the experiment.' (2004:422). The obvious framing which is not discussed by Levitt and List or others, is that lab contexts and experiments are much more unfamiliar to non-western subjects, much more familiar are actual games. Calling an experiment a game, has significant framing entailments (Jackson 2010). In the MacArthur studies there are a number of suggestions that this was the case; the comments on how enjoyable the games were because they were fun (Henrich and Smith 2004: 128), the joking during games (Hill and Gurven 2004: 390) the game explanation by researchers to players in terms of a 'fun game for real money' (Ensminger 2004: 366), the comment made on Ache perceptions of the game as gambling, and observation that one of the most notorious gamblers in the community had contributed very highly to the public good game and was heard to remarking after the game that he had hoped to double his money (Hill and Gurven 2004: 407). High acceptances of low offers were sometimes explained by players in terms of just the bad luck of ending up a responder rather than a proposer (Gurven 2004: 222, and Henrich and Smith 2004: 142), and finally, remarks about game payoffs were seen as windfall money – just as game winning are.

Excessive framing by researchers is seen in the Ache study where players were puzzled about why they would reject an Ultimatum offer, and 'we specifically stated that if the respondent believed the division of the stake was unfair ..he might reject the offer (Hill and Gurven 2004: 391), which tells players how to react, or the Orma case in the labelling of the public goods game as *harambee*.

2.4. Self selection of players

It is clearly important to eliminate distortions and biases arising from the distinctive character of those who volunteer for games, where subjects are not sampled systematically. Volunteers have long been known to have 'more education, higher occupational status, earlier birth position, lower chronological age, higher need for approval and lower authoritarianism than non- volunteers' (Doty and Silverthorne 1975), so it seems curious that self-selection has been acceptable in experimental practice. Levitt and List also remark that the evidence that women are more pro-social than men may be due to their distinctive behaviour in menstrual phases. Women volunteers are more often ovulating (when they are elated and active) than non-volunteers, and they cite evidence that women in experiments on bid the same as men in auctions when they are menstruating, but differently from men when experiencing higher oestrogen levels.

A move away from use of volunteers might seem a ready solution to this area of distortion but even where systematic sampling is used to obtain subjects there are social pressures working against too much certainty of representativeness. In

developing countries the context of poverty, and the use of real money in experimental games, can create enormous desire for inclusion (Jackson 2009) and local gatekeepers facilitating the conduct of experiments, and managing the construction of sample frames, recognise a useful opportunity for patronage – which can easily lead to non-representative participation in games. How to deal with the effects of women's oestrogen cycles is less clear.

2.5. Stakes of the game

The use of real money in experiments is also linked to economist ideas about the external validity of results. Real money produces real behaviour. Thus Vernon Smith writes, 'the laboratory becomes a place where real people earn real money for making real decisions.' (1976:275) A fundamental defence of experimental methods rests on the idea that 'While surveys can provide large and representative data sets that provide statistical power, experiments allow the elicitation of preferences and attitudes in a controlled and incentive compatible way, as participants have to make choices with real money at stake.' (Falk and Heckman 2009: 9).

Experimental economists are generally committed to using real money for several reasons; because 'we feel confident that most people care about it, and that we all want more' (Guala 2005:236); because it is assumed to be universally attractive; and because it is consumed by an individual. Money is also thought to have a positive effect on the veracity of experiments by inducing cognitive exertion (subjects try harder when money is at stake), by increasing motivational focus so that behaviours are easier to interpret, and by emotional triggers - since subjects can behave differently confronted with real money, rather than imagined money (Read 2005).

Levitt and List argue that the larger the stake the more likely it is to dominate moral concerns in experimental behaviour. Benz and Meier (2008) think the effect of the stake on the experiment turns not only on stake size, but also whether the money is earned or obtained in a trivial way. Are these, however, western preoccupations – ie entitlement depends on effort, and more is always better? Anthropologists would approach the use of real money rather differently, and begin with assuming that money is not a uniform, abstract and universal currency but one freighted with symbolism, and meaning, and of differential value to subjects. Money based exchange, is no longer ringfenced and contrasted with the exchange of gifts in 'a 'great divide' between the monetary and pre-monetary worlds' (Bloch and Parry 1989:29), and it cannot be treated as obliterating the rich symbolic and social entailments of other forms of exchange. Money has a concrete materiality, notes of different denominations have character and nicknames (interestingly, often being named after car models), the Bank of Colombia has a major problem with the defacing of banknotes through their use for love notes or drawing on, and Iran too political defacement is common (Guardian 2010). Money has symbolic and magical

meanings and power (Miller 2005), it is evidently scarce in poor communities of the south, and especially so for particular social groups with limited access to monetized relations. Interestingly, the form of money, notes or coins, has also been observed in western experiments to have an impact on game decisions (Smith, 2009, pers. comm.) The value of the stakes was relatively very high in many of the MacArthur studies, eg 10 days wages for Alvard (2004: 418), the scarcity of money commented on as well as the effects of money as the medium of pay-off – Marlowe argues that Hadza food is shared because it cannot be concealed, money is not because it can be hidden (2004: 189).

If money is integrated into other forms of exchange, and has symbolic and cultural character, then we have to consider what meanings and institutions and norms are triggered by the use of money in the games, and therefore what rules might be reflected in offers, acceptances and rejections. The laboratory is not a neutral impersonal context, the money is not an abstract and uniform medium of exchange, and we do not know how the location, the experimenters, the other players and the currency itself was construed by the subjects. I would guess that the high degree of variability in the data from non-western societies derives in part from the highly variable ways in which people struggled to identify the context and expectations.

Guala argues that experimental evidence cannot, alone, bridge the gap between real-world and hypothesis being tested, as, from a different position, does Jackson (2009). Empirical research is needed both before the experiment to ensure the relevant aspects of the real world, and a precise specification, are built into the experimental design, and after the experiment to look for evidence of the same relationships in the real world. Experiments need representative sampling of the relevant population and must also mirror conditions and environments. The trade-off is that experimental results are more reliable when conditions are more artificial, but have more external validity when they are less artificial. Taking external validity seriously implies an important role for empirical research before and after the experiment. '[T]he external validity problem is empirical in character and must be solved by appropriately combining field and laboratory evidence.' (Guala 2005: 160).

3. External validity in the MacArthur studies

In the overview essay on the 15 MacArthur studies, Henrich is confident that play in games has external validity and parallels real life, but reading the individual studies gives a different impression – that this is at best a strained connection, that the parallels are highly selective and speculative, and that the precise ways which social institutions and norms produce particular game behaviours is not clear. Connections made between games behaviours and external reality were quite variable, and more indirect or absent than the overview essay indicates. Some like Ache Indians of

Paraguay are known for their extensive food sharing and cooperative food acquisition (Hill and Gurven 2004: 387), yet their game behaviour shows no relationships between food sharing or food production and cooperation. Others, like Henrich and Smith (2004: 159) talk up the external validity of their study, arguing that the ethnographic data supports their game results as the hunter gathering Machiguenga are individualistic, independent, do not punish or monitor, and do not cooperate - and this accords with their game behaviour of low offers and few rejections of offers.

Claims for external validity are very selective. The non-cooperative gaming of the swidden agricultural Mapuche, whilst very different in their fear of witchcraft and envy, belief in punishment and retribution, is seen as connected to a similar absence of cooperation. They note Mapuche cooperation over religious festivals, but claim this is not 'cued' by the games because they are administered by outsiders, there is an absence of ritual content, and the game is a cashbased transaction. They also mention that livelihood cooperation does exist in female-headed households, but no analysis of play by such subjects is offered. Thus very different societies produce the same kinds of game play because both are selectively represented as somehow essentially non-cooperative, and where cooperation exists, it can be discounted by the suggestion that the framing of the game limits the field of norms to those of interest to the researcher. Such confidence in the precision and certainty over exactly what players thought they were doing in the game seems misplaced.

3.1 Cultural context, game design and the interpretation of evidence:

Guala (2005) argues that empirical research prior to experiments is needed to ensure that there a precise hypothesis formulated for expected behaviour, and that the most salient features of the external reality are built into the game design. It implies that the comparative use of common games in different cultures is very problematic, since salient features of differing cultures cannot be built into a standard game. The MacArthur studies were conducted by researchers with extensive knowledge of their sites, but the use of standard Dictator and Ultimatum Games, designed in western contexts, across all sites for comparability purposes, means that salient features of particular cultures are not built into game design. Further, it is assumed that offers and rejections in games everywhere equate to the same thing – degrees of selfishness, of pro-social behavior, or willingness to punish the breaking of sharing norms, whilst there is evidence that the ways offers and rejections were conceived was very culturally specific in the MacArthur studies. In such circumstances it is difficult to know how the game behavior connects to the external reality, and what behaviours mean.

Rejection levels are shown to reflect a lot more than willingness to punish norm infringements, as the Ultimatum game is supposed to show. Low levels of rejection are seen as possibly connected to reputation fears, and a reluctance to be seen as confrontational – which rejecting a gift would be – for the Ecuador study (Patton 2004: 120). A rejection here amounts to an accusation of meanness and a spiteful act to punish it, thus women are less likely to reject, as are those in weaker coalitions. Here rejection levels may indeed be seen as showing the weakness of a sharing norm, but this is generated by a fear of confrontation, which is altogether different from the rejection levels of western students. In a nother study, Hill and Gurven argue that 'the Ache failed to reject offers because to do so would be a form of serious interpersonal confrontation…[When a person is offered a very small meat share they will not confront the divider but].. instead the recipient grumbles to other members of the social group about the size of the share, thus damaging the reputation of the sharer.' (2004:404-5). Infringement of sharing norms in this contexts is dealt with through mechanisms other than outright rejection – it does not signal the weakness of a sharing norm because subjects are unwilling to punish an infringement *within the game*.

The high offers in PNG (Tracer 2004: 252), compared to other studies, is seen as related to the cultural emphasis on generosity, and the exceptionally high rejection rate to the desire to avoid incurring unwanted obligations by accepting offers – in spite of assurances of anonymity. In this study and in others it is clear that players did not really believe in anonymous play, a situation discussed below. Here, requests to be given things have to be complied with, and people are shunned if they refuse. Unsolicited giving binds individuals together too, with an implied debt, and thus gifts may be refused if the receiver prefers not to be so indebted. Rejections were accompanied by statements such as ''I can take the K3 I received from you [the turn-up payment], I can't take money from someone in the village…' When offered sums above K5 they often seemed genuinely afraid, and on several occasions responders remarked 'no, that's too much.'' (Tracer 2004: 255). Here rejection levels are not connected to willingness to punish sharing norm infringements, as the Ultimatum Game has it, but to avoidance of obligation incurred by acceptance. We also see in this quote that more money is not always better.

External validity requires game design to incorporate salient cultural features from that external reality. If it does not, then the meaning of game behavior is uncertain since it is based on flawed assumptions, in this case about offers, rejections and acceptances, and the extrapolation to external reality becomes questionable.

3.2 Plurality of cultures, norms and meanings:

A further problem arises over the manner in which the external realities of cultures are described in these studies; they characterize cultures in rather monolithic and

generalizing ways with little discussion of internal divisions and contradictions; cultural stereotyping shadows the research. At times particular cultures are spoken of in terms of overarching sharing norms, and at others lower level norms, eg around meat sharing are invoked as evidence of fit with external reality. We have little idea which of the multiple, competing and inconsistent norms at play in even one sphere of life (livelihood production and consumption, religious life, or whatever) are jostling in the minds of players during the games. Thus while McElreath (2004:344) explains differing proposals and rejections in relation to farmer and herder livelihoods, he remarks that, of course, the game may confuse several functionally unrelated elements of social lives – and proposers and responders may be drawing on different norms and thus effectively playing different games.

The ideas which influence game behavior are not only pre-existing norms about sharing, fairness and cooperation, as the experimenters hope, and not only the moral costs and benefits Levitt and List refer to, but include a whole set of diverse ideas about money, goods and gifts, the social meaning of which is only partly about sharing as such.

A reality check on the interpretations and extrapolations of game results on social preferences is sounded by Benz and Meier who comment on the fact that, outside of the lab, the same people seem to show very different social preferences depending on the context, and they conclude that 'Individuals behaviour seems to be extremely situationally dependent and very hard to generalise – either because there are no cross-situational traits or because pro-social preferences are triggered differently in various settings. ..[P]eople's behaviour correlates only weakly between various situations - independent of whether the decision situations are inside or outside the lab.' (2008: 280). Recently experiments have found that subject who chose to buy an ethical product in a supermarket were more selfish in the subsequent Dictator game, and more likely to lie and steal (Mazar and Chen-Bo 2010). They conclude that 'people do not make decisions in a vacuum; their decisions are embedded in a history of behaviours' (2010:10); and that not only can one act, which creates a halo effect, licence an opposite behaviour (see also Monin and Miller 2001), but this may spread across domains, and thus a general sense of moral self influences, and licenses, pro- and anti social behaviour. A certain quota of pro-social behaviour soothes a conscience, and is felt to legitimate selfishness.

The problem of the consistency of results obtained from the same players in different games, is commented on by Gurven (2004: 222) who asks, 'If economic games capture real-life preferences, then how can no relationship exist between Ultimatum Game offers and Public Goods game contributions made by the same individuals?'.

3.3 Expectations about norms and behavior in experimental methods:

Experimentalists assume that what players do in the experiment reveals the strength of norms of sharing and fairness, and thus how they assume norms produce actions is important for the interpretations given to game evidence. In game theory, norms are the conventions, the rules of thumb, which indicate appropriate behavior, but people only observe them if they are optimal for them. Experimental games approach norms in a clearly functionalist way – they work to coordinate action, they produce stability through general adherence, and they serve various institutions, eg trust is functional to markets. Game behaviour is expected to display the norms which inform everyday life, and norms are the means by which institutions and social structures reproduce themselves.

The universe of norms however is large and complex, with many contradictory norms, and alongside the hegemonic norms of how things ought to be done sit 'muted' social structures (Jacobson-Wilding 1999) and subaltern norms. And they exist at many levels, from minor social conventions to life and death matters, and carry very different sanctions. Norms are 'contingent claims that have to be 'made to count' through the effective mobilization of sanctions in the contexts of actual encounters' (Giddens 1983: 30). However, actual encounters, theorized by Goffman, show social interaction in specific situations to be rather loosely geared to social structures and norms. For him, 'Social situations can consolidate structural lines or loosen them' (1983), and a *bricolage* of norms may be informing player's choices. For game theorists norms govern behaviour as rules of thumb, unless over-ridden by rational self-interest, but to a social theorist of a more actor-oriented ilk, norms present an array of more or less sanctioned, sometimes ambiguous and contradictory, often unstable, social constraints and discursive resources with considerable room for interpretation by actors. Neither Goffman (1961), Giddens (1984) or Long (2001) would have expected game play to connect straightforwardly or easily to 'real life' behaviours. And how does the researcher know the difference between a player blindly following a norm, operating a counter-norm, inverting a norm, or deciding that the context suggests a domain of completely different norms, such as a real game does?

The idea of external validity assumes a direct mirroring of behaviour in games and in 'real life'. But the labeling and framing of the encounter as a 'game' may suggest other relations to norms within the encounter. The whole point of a real game is that it is not real life. Games might be expected to be engrossing fun because they have uncertain outcomes, opportunities to display personal qualities valued in the wider world, to take risks and gamble character. What is therefore at stake in experimental games is always much more than the money stake. Far from simply acting in line with hegemonic sharing norms, a player engaged in 'character contests' may gain from counter-normative behaviour. A game is partly bounded from the external

world; it is no longer fun if it is too close to reality or too far from it. 'It seems .. that in games, and similar activities, disguises must be provided which check, but do not stop, the flow of socially significant matters into the encounter' (Goffman 1961: 73).

One of the MacArthur authors does briefly suggest an inversion of the 'games reflecting real-life' idea; Marlowe on the Hadza, finds a paradigmatically egalitarian society behaving in unegalitarian ways – their modal offer was 20% in the UG compared to the 50% among western students (2004: 175). Since everyday life involves a great deal of sharing Marlowe suggests that this behaviour is the reverse of games reflecting reality, for 'the more frequently one must share, the more weary one grows of it, and the more one looks for any opportunity to escape it.' (2004:187) Games could be just that – an act of resistance or an amusing diversion standing outside everyday rules.

The arguments made for external validity seem to me strained and inconsistent, with little consideration of how far the terms of the discussion map from one culture to another, how plural is the universe of norms available for any particular context and how creative individuals are in their deployment of these norms They also have a limited specification of how norms are expected to relate to game behaviours – are they mirrored or inverted, determining of action or simply a backdrop to agentic manipulations? The terms of the discussion are not, in themselves, scrutinised cross-culturally, ie what is generosity or altruism or selfishness or cooperation in other societies? And, critically, how do they see money, gifts and exchanges of various kinds, and how does this differ from western ideas? What are games in the cultures where these studies took place? The experiments may be rather casually labelled 'games' but this is a clear signal to players. Finally, there is too little reflexivity in relation to how the game findings are patterned by the experimental exchange itself (Jackson 2009). There are several studies which hint at the influence of the researchers and their assistants in cueing the exercise of particular norms and thus the apparent application of sharing norms from the social institutions prefigured. The research assistants in the Orma study immediately saw the games as related to *harambee* (local self-help collective projects), and the players unsurprisingly came to this view too.

Conclusion

Many of the authors of the component studies of the MacArthur collection are less ready to generalize findings or to accept game results as possessing external reality as readily as the overview essay does.

Interpreting game behavior in terms of fairness norms was clearly problematic: fear of punishment or confrontation rather than fairness norms drove offer and acceptance levels, acceptance of low offers was connected to groupishness for the Tsimane (this retains the pay-off in the group rather than seeing it retained by the researcher) (Gurven 2004: 221); and offer rejections were seen as implying one is not a dupe (Alvard 2004: 425). Criticisms of the method included the lack of consistency of play across games, producing widely varied results for different games in the same locations, the difficulty of glossing a society as cooperative, or not, when there are elements of both in most, knowing which cooperative norms are exercised in games, the problem of knowing whether play is reflecting the norms of everyday livelihood cooperation (as most seek to find) or simply etiquette or manners (Gurven 2004: 226). The most detailed account of the multiple misunderstandings and confusions about the experimental game is in Gil-White's account of his Mongolian research, from which he concludes that 'My methods, careful as they were, could not anticipate the great cultural gulf separating me from my respondents, for I could not have imagined some of the hypotheses they made concerning the object of the game.' (2004: 273). The legion ambiguities deriving from the game design, and hypotheses, make interpretation of the meaning of game results challenging, and leaves an extremely wobbly foundation for the claims to external validity.

Questions about external validity were also raised by Hill and Gurven, finding no association with real life behaviours (2004:408), and Gurven for the Tsimane who finds it hard to explain high variability between villages, and cautions against 'attempting to explain game results by fitting just-so anecdotal stories that capture key cultural traits' (2004:225), while the generalization made about market integration and fairness norms finds very patchy support in individual analyses.

Experimental games confront serious difficulties in making claims which rest on this evidence alone. What is required is a triangulation of game evidence with other sources, much better knowledge of cultural contexts and more carefully specified ideas of what exactly games are testing, greater researcher reflexivity, and a systematic debriefing after games to ascertain what frames and ideas informed play. The potential value of standard games designed in the west in other cultural contexts is questionable, and at the very least requires a hugely greater attention to local perceptions of the encounter.

References

Benz M and Meier S (2008) 'Do people behave in experiments as in the field? Evidence from donations' *Experimental economics* 11, pp268-281

Bloch M and Parry J (1989) (eds) *Money and the morality of exchange* CUP

Charness G and Kuhn P (2010) 'Lab Labor: What Can Labor Economists Learn from the Lab? National Bureau of Economic Research, Working Paper No. w15913

Chen Y, Katuscak P and Ozdenoren E (2005) 'Why can't a woman bid more like a man?' Dept of Economics, University of Michigan, Working Paper

Croson R and Gneezy U (2009) Gender Differences in Preferences *Journal of Economic Literature*, Volume 47, Number 2, pp. 448-474

Doty R and Silverthorne C (1975) 'Influence of menstrual cycle on volunteering behaviour' *Nature* 254 139-40

Eckel C (2008) The gender gap: using the lab as a window on the market' in Bettio F and Verashchagina A (eds) *Frontiers in the economics of gender* Routledge Siena Studies in Political Economy, 223-242

Falk A and Heckman J (2009) 'Lab experiments are a major source of knowledge in the social sciences' *Science* 326, 535-538

Giddens A (1984) *The constitution of society* Polity Press

Giddens A (1990) *The consequences of modernity* Stanford University Press

Goffman E (1961) Encounters: two studies in the sociology of interaction Bobbs Merrill: Indianapolis

Guala F (2005) *The methodology of experimental economics* CUP

Guardian, April 10, 2010 'Colombia tries to stop people drawing money'

Haley K and Fessler D (2005) 'Nobody's watching? Subtle cues affect generosity in an anonymous economic game *Evolution and Human Behaviour* 26 245-256

Hoffman E, McCabe K, Shachat K, Smith V (1994) 'Preference, property rights and anonymity in bargaining games' *Games and Economic Behaviour* 7(3) 346-80

Jackson C (2009) 'Researching the researched: gender, reflexivity and actor-orientation in an experimental game' *European Journal of Development Research*_Vol 21, 772–791

Jackson C (2010) 'Research with experimental games: questioning practice and interpretation' *Progress in Development Studies* (forthcoming)

Jacobson-Wilding A (1999) 'Chaos and creativity: the transformative symbolism of fused categories' in Moore H, Sanders T and Kaare B (eds) *Those who play with fire: gender fertility and transformation in east and southern Africa* LSE Monographs in Social Anthropology, Athlone Press, pp283-293

Levin I, Schneider S, Gaeth G (1998) 'All frames are not created equal: a typology and critical analysis of framing effects' Vol 76(2) 149-188

Levitt S and List J (2007) 'What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?' *The Journal of Economic Perspectives*, Vol. 21, No. 2 pp. 153-174

Long N (2001) *Development sociology: actor perspectives* Taylor Francis.

Mazar N and Chen-Bo Z (2010) 'Do green products make us better people?' *Psychological Science*

Miller D (2005) (ed) *Materiality* Duke University Press

Monin B and Miller D (2001) 'Moral credentials and the expression of prejudice' *Journal of Personality and Social Psychology* 81(1) 33-43

Read D (2005) Monetary incentives, what are they good for? *Journal of Economic Methodology*, Vol 12, (2) pages 265 - 276

Smith V (1976) 'Experimental Economics: Induced Value Theory' *The American Economic Review*, Vol. 66, No. 2, pp. 274-279

Tversky and Kahneman (1981) 'The framing of decisions and the psychology of choice' *Science* 30, Vol. 211. no. 4481, pp. 453 - 458