

Adam Smith's Fair and Impartial Spectator: Fairness was About Fair-Play Rules for Society, a Pre-Condition for Wealth Creation

Vernon L. Smith
Chapman University

In his first book, Adam Smith used the metaphor of the fair and impartial spectator to articulate his model of the maturation process whereby people learn to follow general rules (norms) that take into account the human sentiments of gratitude and resentment in others. Through the impartial spectator, human action is governed by self-command. I will show why, in Smith's model, our self-interested nature is a key part of the process of learning to act in sympathetic other-regarding ways in our close knit communities; how the rules of propriety naturally inform property rights in the civil order, and thus to set the stage for explaining the wealth creation process articulated in his widely celebrated second book.

[T]here are indeed some universal moral norms and values, but to think that 'fairness' is among them is an Anglocentric illusion. - Anna Wierzbicka (2006, p. 162)

Although Adam Smith referred to the "impartial spectator" over 70 times in his first book, *The Theory of Moral Sentiments* (TMS)(Smith, 1790), he spoke only once of the "fair and impartial spectator": "We endeavor to examine our own conduct as we imagine any other fair and impartial spectator would examine it" (Smith, 1790, p. 110). This form of the metaphor best enables us to understand Smith's conception of the maturation process wherein we become socialized by gradually modifying our behavior to follow general rules that meet with the approval, and to avoid the disapproval of our neighbors.

I will use a propositional style of discourse to articulate and discuss Smith's model of human sociality, and the central role of the metaphor of the impartial spectator, beginning with some propositions that provide background axioms and principles. The power of Smith's work is that it accommodates the observed tendency for humans to

regard others in their more intimate groupings, explains the emergence of property as it occurs in the civil order of government, and accounts seamlessly for the prominence of self-interested action in impersonal markets, leading to specialization and innovation, the cause of the wealth of nations. We are strapped in finding a modern equal to Smith's grand accounting for the deep meaning he extracted from carefully observing the diversity of human conduct.

Proposition 1: In Smith's model, learning to become social is not about altering our self-interested or self-loving nature, but rather incorporating that nature into a theory of the emergence of socializing rules through processes of cultural consent.

In modern language, each person is characterized by strictly increasing individual utility functions defined on their own valued outcomes, say $U(\text{own})$. If we think of an outcome as always having a monetary equivalent, then utility is strictly increasing in monetary amount. This non-satiation axiom (that for each of us more is better and less is worse) is common knowledge. Thus,

Every man is, no doubt, by nature, first and principally recommended to his own care; and as he is fitter to take care of himself than of any other person, it is fit and right that it should be so (Smith, 1790, p. 82).

and, [. . .] every animal was by nature...endowed with the principle of self-love [. . .] (Smith, 1790, p. 272).

We are not ready to suspect any person of being defective in selfishness. This is by no means the weak side of human nature, or the failing of which we are apt to be suspicious. Carelessness and want of economy are universally disapproved of, not, however, as proceeding from a want of benevolence, but from a want of the proper attention to the objects of self-interest (Smith, 1790, p. 304).

Generations of modern economists were indoctrinated with a thought process in which every action maps into an outcome and thence into preference and, implicitly, in which this mapping can be reversed via individual maximization. Where action is other-regarding instead of strictly self-regarding, it is tempting to simply change the arguments of the utility function to include other as well as one's own outcomes. Smith's model, however, is in no sense based on the hypothesis that individual preferences, redefined as including own and other outcomes, are the source of the concern people have for others. This is explicitly stated and defended in the above quotations, and implicitly assumed in the key propositions on beneficence (7 and 8) and justice (9 and 10) described below.

Indeed, the opposite is implied; common knowledge of self-love is essential if we (through our impartial spectators) are to make appropriate judgments concerning proper conduct in human social relations. Contrary to neoclassical and modern economics, for Smith, self-love did not imply that individuals would in all, or perhaps even in most, circumstances choose actions to maximize the utility of own outcome,

U(*own*). Rather, it is through knowing that all in an interactive community (extended family, neighbors, associates) are self-interested that we know that a given action is hurtful to anyone who receives less, and beneficial to anyone who receives more. In Smith's model of sympathetic fellow-feeling, common knowledge of how alternative actions hurt or benefit others as well as oneself, provides the foundation whereby people learn to follow rules that are appropriately other-regarding, that properly take into account the feelings (the gratitude and resentment) of others.

If you do not know what hurts or benefits others, you cannot know how to modify your actions in order to live harmoniously with others.

The neoclassical and modern error is to apply the Max U(*own*) calculus to all decisions, regardless of circumstances, and without regard for the pattern of benefits and potential hurt in our more intimate groupings where enforcement was, and always had been, endogenous. The behavioral economic "social preference" error is to replace that description with a just-so utility function of the form U(*own, other*), in effect rescuing Max U. The methodological error is to focus on outcomes instead of process. This is why Smith also gave us clear thinking on liberty as the first principle of human society. He believed that liberty was the source of sustainable social and economic harmony.

Proposition 2: Social motivation arises from the individual's desire for praise and praise-worthiness and the desire to avoid blame and blame-worthiness, which serve the fundamental values of propriety and harmony in the evolution of local order from local rules.

[T]o feel much for others and little for ourselves, that to restrain our selfish, and to indulge our benevolent affections, constitutes the perfection of human nature; and can alone produce among mankind that harmony of sentiments and passions in which consists their whole grace and propriety (Smith, 1790, p. 25).

Man desires, not only praise, but praise-worthiness; or to be that thing which, though it should be praised by nobody, is, however, the natural and proper object of praise. He dreads, not only blame, but blame-worthiness; or to be that thing which, though it should be blamed by nobody, is, however, the natural and proper object of blame (Smith, 1790, p. 113-114).

Hence, the motivation for action in our more intimate groupings is not itself utilitarian; rather Smith models the process whereby we modify self-interested choices in the light of learning what other people will go along with. Praise and praise-worthiness are means for describing that social approval, but the resulting approved conventions require each to know the pattern of hurts and benefits resulting from an action. Since all are self-interested, we can judge who is hurt or benefits from an action and integrate that essential knowledge into our learning of rules in which our actions are praised/praise-worthy or are not blamed/blame-worthy. Thus, other-regarding behavior does not derive from any other-regarding outcome utility, U(*own, other*), but rather is the result of U(*own*) as an input to our socialization. Although the new behavioral utilitarian would modify Smith's model by assigning everyone an

individual social preference function, this path was not followed until after economic thinking was shaken by the rejection of neo-classicalism in small group experiments, especially two-person trust and other games in the 1980s and 1990s. The work of Berg, Dickhaut, and McCabe (1995) was the key paper that initiated a large subsequent literature (Smith 2008). By then Adam Smith's work had over two centuries of priority, and deserved careful examination of how he was able to model small group behavior within the framework of self-interested individuals. Moreover, in impersonal markets we also rely on *U(own)* in choosing to take action. Wilson (2010, p. 78-81) contrasted how the word "preference" is used to interpret market decision with how it applies to social interactions. There is no need to model the individual as a divided self; rather, we can model one self-interested individual in imperfect self-command of his local relationships while simultaneously responding to the external order of prices in markets. However, economic and social policy is threatened by human failure to understand that the rules of the local order cannot be applied to those of the extended order, or vice versa, without damage to the one or the other (e.g., Hayek, 1988).

Proposition 3: The process of learning to be sociable—maturation—is to learn propriety.

Smith uses an ingenious mental experiment — his *soziale Gedankenexperiment* — to articulate his conception of the socializing process. We are asked to imagine an individual growing up without any communication with another human being. For Smith, such a solitary individual "could no more think of the propriety or demerit of his own sentiments and conduct, of the beauty or deformity of his own mind, than of the beauty or deformity of his own face" (Smith, 1790, p. 110, 192-193). A solitary person can see none of these things in the absence of a social mirror. Raise him in society, and that mirror is supplied in the form of the "countenance and behavior" of all who he lives with, who never fail to express their sense of the propriety or impropriety of his actions (Smith, 1790, p. 110). From this experience each individual is able to internalize a view of the appropriateness of his own conduct and gradually acquire personhood. For Smith, "mind" is a social creation, whether it involves our conduct in the choice of context-dependent action or our perceptions of facial or body beauty. There is no individual psychology separate and distinct from social psychology. Psychology in this sense must begin with principles acquired from our human sociability.

Proposition 4: The concept of 'fairness' lives in rule space, and it corresponds to the sports metaphor of fair play in which people are motivated to choose actions that avoid committing fouls.

Hence, Smith uses *fair* in its 18th century meaning that was, and is, a unique English word. As observed by the distinguished and influential linguist, Anna Wierzbicka (2006):

The ubiquity of the words *fair* and *unfair* in modern English discourse, across a wide range of registers, is all the more remarkable given that these words have no

equivalents in other European languages (let alone non-European ones) and are thoroughly untranslatable (p. 141).

And again,

‘fairness’ is a uniquely Anglo concept, without equivalents in other languages, except, as for example in German, as a loan from English (*das is nicht fair*, “that’s not fair”). At the same time, in Anglo culture this concept is so central that many speakers of English imagine it must be universal, perhaps even innate ... there are indeed some universal moral norms and values, but to think that ‘fairness’ is among them is an Anglocentric illusion (Wierzbicka, 2006, p. 160-162).

And in Polish it is to *nie fair* where Wierzbicka (2006) reported resisting this word loan when her bilingual daughters, contrary to her own native language and cultural experience, began thinking in terms of the English word. Her final summary applies without modification to my representation of Smith and his concept of the fair and impartial spectator in this paper:

In a way, sport — especially team sport — provides a perfect model for ‘fair’ interaction because the emphasis is on rules and procedures, which are blind to the individual players’ interests and which everyone voluntarily accepts (Wierzbicka, 2006, p. 166).

I want to elaborate on Wierzbicka’s team sport metaphor in relation to Smith’s idea of fair interaction. When teams are playing other teams, their play must accord with the rules, with no fouls allowed by the referee who monitors and levies appropriate penalties under the rules of team vs team conference play. These are analogous to “property rights” in the economy, which govern the interaction of economic agents in their economic world. Within each team there are informal rules by mutual consent that support coordination and leverages the group’s action in competition. These internal rules of “fair-play” may be broken by an individual tempted to leverage his recognition by favoring himself over assistance to another that would improve coordination effort and team score over the individual’s score. These are analogous to “propriety rights” in *The Theory of Moral Sentiments*. When these rules are broken, dissention sets in and team effort is disharmonious; outwardly it is inefficient, but Smith is examining its deeper cause in human sociality.

Proposition 5: The metaphor of the fair and impartial spectator defines the processes whereby we first judge the conduct and character of our neighbors, then devolve or pass judgment concerning the conduct and character of ourselves.

These two parts of the judgment process are paraphrases of the subtitle of TMS from the fourth edition, which appeared in 1774 (e.g., Raphael & Macfie, 1976, p. 40). The judgment concerning our own conduct gradually takes the form of self-command which evolves from our two-state experience with failures of character in the marginal

moment (the man of yesterday) that are reconsidered in the cooler light of subsequent reevaluation (the man of today). Thus there are two occasions wherein we are afforded the opportunity to view our conduct from the perspective of the impartial spectator. The first is at the time we are poised to act. The second occurs after having acted. In both cases, our sentiment is quite partial, but it is the most partial when it is important that it be impartial. At the time of action, the passion of the moment interferes with an impartial evaluation. Although afterward, the prompting circumstances and passion allow a cooler impartial judgment, too often — in comparison with the heat of the moment — the consequence seems unimportant, and except for vain regret, we may fail to secure ourselves from like errors in the future (Smith, 1790).

This self-deceit, this fatal weakness of mankind, is the source of half the disorders of human life. If we saw ourselves in the light in which others see us, or in which they would see us if they knew all, a reformation would generally be unavoidable. We could not otherwise endure the sight (Smith, 1790, p. 158-159).

Proposition 6: We are rescued, however, from the frailties of our conscious judgments by our stronger tendency toward ingrained rule-following conduct — general rules that map particular circumstances into actions that inspire the gratitude, and avoid the resentment, of others.

Fortunately for our species, nature has not entirely abandoned us to the delusions of self-deceit triggered by our self-love. From our earliest exposure to the conduct of others, we gradually become attuned to general rules that constitute acceptable “fit and proper” actions sensitive to the context in which they take place (Smith, 1790, p. 159).

According to Smith, our conduct takes key categorical forms that I will summarize in the next four propositions. The first two govern beneficent actions; the second two concern hurtful actions, and they encapsulate Smith’s theory of justice and property rights.

Proposition 7: Intentionally beneficent actions alone deserve reward because of the gratitude invoked in others (Smith, 1790, p. 78).

This proposition provides the emotional foundations of reciprocity, a universal concept requiring an explanation. Thus for Smith, reciprocity is not a satisfactory explanation of the choice outcomes observed in trust games, as in McCabe, Rassenti, and Smith (1996). Rather, ‘reciprocity’ is an un-modelled name for the result we observe, and Smith sought a deeper explanation (for a careful discussion of these issues see Wilson, 2008). Our beneficence is most naturally directed to those whose beneficence we have already experienced, and therefore kindness begets kindness (Smith, 1790, p. 225). In repeat interaction with our associates, reputational gains from sociability yield human betterment, and “tend to unite men in society, to humanity, kindness, natural affection, friendship, esteem” (Smith, 1790, p. 243). This phenomenon is captured in the modern phrase ‘I owe you one’ — common across many languages — in which the beneficence of another is acknowledged by an implicit obligation to do a future favor in return. The

debt is discharged by an ‘in kind’ transfer (i.e., in the same way, with something similar). One cannot resist interpreting the exchange as de facto ‘in kind-ness.’

Proposition 8: The want of beneficence cannot provoke resentment and punishment, because beneficence is freely given and cannot be extorted (Smith, 1790, p. 78).

These two propositions have been tested in the context of extensive form trust games (Smith & Wilson, 2014, 2016). Under anonymous pairing, the traditional game-theoretic analysis predicts no cooperation. However, in accordance with Proposition 7, half or more of first-movers beneficently offer cooperation, and two-thirds of their paired counterparts eschew the more lucrative opportunity to defect, instead rewarding the first mover by choosing the cooperative outcome. Proposition 7 retroactively predicts the findings in early trust games better than neoclassical economic analysis. In new experiments, the test of Proposition 7 is replicated, and a modification of the same game is used to test Proposition 8. If first movers choose not to offer cooperation, play passes to second movers who are provided a costly option to punish their paired counterpart for failing to offer cooperation. None choose this option. Implicitly, the second mover’s response in these experiments freely acknowledges the right of the first not to act beneficently by offering cooperation.

Proposition 9: Intentionally hurtful actions alone deserve punishment because of the resentment invoked in others. The greater the hurt, the higher the resentment, and, in proportion, the greater the punishment (Smith, 1790, p. 78, 83-84).

This proposition is the foundation of Smith’s theory of property rights. Our human impulse is to punish intentional actions of a hurtful nature:

Resentment seems to have been given us by nature for defence, and for defence only. It is the safeguard of justice and the security of innocence. It prompts us to beat off the mischief which is attempted to be done to us, and to retaliate that which is already done; that the offender may be made to repent of his injustice, and that others, through fear of the like punishment, may be terrified from being guilty of the like offence (Smith, 1790, p. 79).

Accordingly, in the civil order of government we find that murder, the greatest evil, commands the greatest punishment; theft and robbery, which deprive us of our lawful possessions, command a greater punishment than the violation of contract which merely frustrates our expectation of gain (Smith, 1790). (See Proposition 11 for further discussion and explanation of the differential penalties for theft or robbery versus violation of contract).

Justice for Smith is a negative virtue that results in a large residue of allowable actions after using proportioned punishment to limit specified hurtful actions of injustice. In Smith’s conception, we do not set our sights on a positive ideal of justice — an abstract, slippery and vaguely defined state. Rather, we address ourselves to

specific acts of injustice where — as I interpret Smith — we are likely to find common agreement because of our common experience of the circumstances, nature, and extent of the hurt. Eliminate these infractions one by one, and in this evolutionary process we gradually produce a more just society, but always within a framework of freedom to act and explore all options not specifically interpreted as unjust. This model is severely challenged today as university campuses are beset by new conflicts between the traditional right of free speech and expression, and the demand for new rights in which various social group identities (atheists, women, sexual orientation, ethnicity, race, etc.) want protection in the classroom and in campus commons against offensive speech or conduct or they seek group space (restroom) assignment, or separate but equal activity space; all in the interest of feeling safe and comfortable. The consequence is wide disagreement on what the rules should be.

At this juncture in the discussion it is natural to ask which of these two sentiments — beneficence or justice — is the more essential to human society. On this, Smith leaves us with no doubt as to his views. We are informed that society will certainly flourish if it is bound by a common bond of gratitude, friendship, and esteem, but where these conditions do not exist, society, though reduced in happiness, may nevertheless not be dissolved. For society can subsist merely from a common sense of its usefulness, as with a group of merchants, and be supported by a mercenary exchange of good offices according to an agreed valuation.

This contrast, between a more intimate social in-group and one bound by a general recognition of the usefulness of association, is illustrated in the evolution of an experimental economy studied by Kimbrough, Smith, and Wilson (2008). The economy consisted of three dispersed villages, each consisting of four houses and their associated fields. Each village produced two out of three world products, and each individual village member received private utility from all three products. Hence, each village had to trade with at least one other village to fully prosper. Two members in each village were empowered to travel to a common “merchant” area where trade could occur, then return to their home villages. In the course of the experiment each village attained a degree of closeness never matched by the merchant area. The village chat rooms were alive with the use of “we,” whereas “the interactions in the merchant meeting area are noticeably more impersonal than those in the villages” (Kimbrough et al., p. 1025).

Society, however, cannot subsist among those who are at all times ready to “hurt and injure one another” (Smith, 1790, p. 86). Hence, beneficence is less critical to a society's existence than justice. Although society may subsist in the absence of beneficence, it will soon be destroyed by rampant injustice. Smith's oppositions to slavery, mercantilism, imperialism, colonialism, and taxation without representation were firmly rooted in his theory of socioeconomic development. Thus (Smith, 1790, p. 81-82),

Proposition 10: Choosing to forgo actions of a hurtful nature does not merit reward.

While in the civil order of law we punish infractions of justice, we do not reward people for obeying the law. There is no reward for stopping at a red light or for leaving your neighbor undisturbed. These are your duty, and call for no explicit rewards,

though in following the law we hope that others will do likewise and all benefit.

Proposition 11: There is an asymmetry between gains and losses: “We suffer more when we fall from a better to a worse situation, than we ever enjoy when we rise from a worse to a better” (Smith, 1790, p. 213).

Note that Smith’s fundamental concept of the asymmetry between gains and losses is a modern idea, rediscovered in experimental psychology, and an important element in the recognition of Daniel Kahneman (2003) for the 2002 Nobel Prize in Economics. For Smith, the concept was central not only to understanding aspects of human action, but also in how it informed the content of differential punishment for loss from theft and robbery versus loss from contractual promises. The asymmetry between gains and losses essentially follows from an asymmetry between joy and sorrow. Most people, reasonably situated and not destitute, can rise above that state, but little can be added to that state in comparison with what can be taken from it: “Adversity, on this account, necessarily depresses the mind of the sufferer much more below its natural state, than prosperity can elevate him above it” (Smith, 1790, p. 45). Moreover, this asymmetry is not only a private, or utilitarian, experience:

It is averse to expose our health, our fortune, our rank, or reputation, to any sort of hazard. It is rather cautious than enterprising, and more anxious to preserve the advantages which we already possess, than forward to prompt us to the acquisition of still greater advantages. The methods of improving our fortune, which it principally recommends to us, are those which expose to no loss or hazard (Smith, 1790, p. 213).

Smith (1790) uses Proposition 11 to provide us with an explanation of why punishment is more severe for theft and robbery, a criminal offense, than for violation of contract, which is a civil offense.

If Smith’s *Wealth of Nations* (1776) is read as a sequel to his earlier work, the continuity in his thought is compelling, and it contrasts sharply with post-neoclassical economic thought in the 20th century. For Smith, economic development is the next great step in a culture that has evolved rules of fair play and is accustomed to well-practiced social interaction; trade comes from the same sociability, and thus begins his second book.

By permission, this article is a revised and rewritten version of Vernon L. Smith, “The Fair and Impartial Spectator,” *Economic Journal Watch*, 13(2), May 2016.

References

- Berg, J., Dickhaut, J., & McCabe, K. A. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122-142.
- Hayek, F. A. (1988). *The fatal conceit: The errors of socialism*. Chicago: University of Chicago Press.

- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5), 1449-1475.
- Kimbrough, E., Smith, V. L., & Wilson, B. J. (2008). Historical property rights, sociality, and the emergence of impersonal exchange in long-distance trade. *American Economic Review*, 98(3), 1009-1039.
- McCabe, K. A., Rassenti, S. J., & Smith, V. L. (1996). Game theory and reciprocity in some extensive form experimental games. *Proceedings of the National Academy of Arts and Sciences*, 93, 13421-13428.
- Raphael, D. D., & Macfie, A. L. (1976). *Introduction to the theory of moral sentiments by Adam Smith*. Oxford: Oxford University Press.
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. Oxford: Oxford University Press.
- Smith, A. (1790, 1976). *The theory of moral sentiments*. Oxford: Oxford University Press.
- Smith, V. L. (2008). *Rationality in economics: Constructivist and ecological forms*. New York: Cambridge University Press.
- Smith, V. L., & Wilson, B. J. (2014). Fair and impartial spectators in experimental economic behavior. *Review of Behavioral Economics*, 1(1-2), 1-26.
- Smith, V. L., & Wilson, B. J. (2016). Sentiments, conduct, and trust in the laboratory. *Social Philosophy and Policy*, 33(1).
- Wierzbicka, A. (2006). *English: Meaning and culture*. Oxford: Oxford University Press.
- Wilson, B. J. (2008). Language games of reciprocity. *Journal of Economic Behavior & Organization*, 68, 365-377.
- Wilson, B. J. (2010). Social preferences aren't preferences. *Journal of Economic Behavior & Organization*, 73, 77-82.