

# The Coherence Model of Preference and Belief Formation

Sun-Ki Chai\*

\* sunki@hawaii.edu, University of Hawai'i, Honolulu HI

## 1 Introduction

Rational choice is the dominant approach in the social sciences to modeling individual and collective behavior. Recently, however, a great deal of criticism has been directed at it, much of from practitioners of the approach itself. Perhaps the strongest criticism has been directed at rational choice's inadequate modeling of preferences and beliefs. While a large literature has arisen to examine alternative assumptions about preferences and beliefs, these new works typically limit themselves to a fairly narrow scope or leave unspecified parameters, creating the impression that one must trade away the strengths of the conventional approach in order to attain greater realism. However, this is not necessarily the case. The coherence model presented here accounts for preferences and beliefs in a way that is generalizable across the full range of contexts where conventional rational choice models can be applied and is compatible with assumptions of rational optimization. It will be based on the assumption that individuals will adjust preferences and beliefs to minimize expected regret, where expected regret is defined as the difference in perceived expected utility between each action taken and the action retrospectively viewed as optimal. Coherence is defined as attaining a state of zero expected regret is referred to as coherence, and is seen as the ultimate "meta-goal" of individuals engaging in "choosing" their preferences and beliefs with a specified set of reality constraints. This model designed to integrate a range of empirical findings and theories about the construction of self from a variety of social science disciplines.

The goal of the coherence model is to provide a general model of preference and belief formation, one that is consistent with a rational optimization model of action characteristic of the rational choice approach. It addresses a wide range of criticisms that have been made against the conventional rational choice approach, particularly the accusation that its view of preferences and beliefs is unrealistic and leads to inaccurate or indeterminate predictions. Actually, there are two versions of the rational choice approach that can be recognized. One is "thin" rationality, in which the nature of preferences and beliefs is left unspecified, and the other is "thick" rationality, in which preferences are seen as self-regarding, materialistic, and isomorphic across all individuals; and beliefs are seen as being completely determined by observation and logical inference from that observation (Chai 2001, 5-8). In practice the thick version is

what is generally associated with conventional application of the rational choice approach, since the thinner version is not capable of making predictions on its own.

However, the conventional approach has been attacked from a number of directions as being unrealistic as a portrayal of human nature, as of limited use in making predictions outside a narrow realm of economic behavior. It is clear that one way to improve the predictive accuracy of the model would be to generate new models of preference and/or beliefs. Much of the recent activity has taken place in the growing fields behavioral economics, where the concept of *social preferences* has taken hold as an antidote to the conventional self-regarding preferences (Rabin 1993; Bolton and Ockenfels 2000, Fehr and Schmidt 1999; Andreoni and Miller 2002; Charness and Rabin 2002), but similar activity is also taking place among rational choice practitioners in political science and sociology (for review, see Chai 1997). However, these approaches purposely generally posit an alternative set of uniform preferences and do not attempt to predict the variations in preferences that exist between individuals and groups, and within individuals and groups over time.

The coherence model is aimed at addressing just these sorts of variations by presenting a dynamic model of individual preference and belief formation. Moreover, because the process through which preferences and beliefs are formed are taken as inherently social, involving both collective action and communication, the model is also intended to be one which can explain the generation of social identity and culture as collectively held attitudes.

## 2 Specifying the Model

*Regret* for a single action, given a particular state of the world, is the difference between the maximal utility possible in that state and the utility provided given the chosen action. *Expected regret* for a single action is the regret for each possible state of the world multiplied by the perceived probability of that state's occurrence.

The simplest example of how expected regret might be calculated is in a dichotomous choice set with alternative actions  $a$  and  $b$ , and two possible states of nature  $s_1$  and  $s_2$ , which are perceived to have probabilities  $p_1$  and  $p_2$ . This creates possible outcomes  $U_{a1}$ ,  $U_{a2}$ ,  $U_{b1}$  and  $U_{b2}$ , where  $U_{a1}$  stands for the utility resulting from action  $a$  in state of the world  $s_1$ , and so forth. The expected regret for choosing action  $a$  would then be  $d = p_1 (\max(U_{a1}, U_{b1}) - U_{a1}) + p_2 (\max(U_{a2}, U_{b2}) - U_{a2})$ .

For more complex choice sets with choices  $a_1 \dots a_n$ , possible states of nature  $s_1 \dots s_k$ , perceived probabilities  $p_1 \dots p_k$  for each state, and where utilities  $U_{ij}$  stands for the utility resulting from action  $a_i$  given state of the world  $s_j$ , and  $U_j^* = \max(U_{1j}, \dots, U_{nj})$ , expected regret for choosing action  $a_1$  will be  $\sum_{j=1}^k p_j (U_j^* - U_{1j})$ . Where outcomes are continuously distributed along a single dimension and  $\phi_s$  is the probability density

function for  $s$ , this will be  $d = \int_s (U_j^* - U_{1j}) \phi_s(s) ds$ . It is clear from the definition that  $d$  will always be non-negative. Based on this definition, expected regret can be calculated whenever individuals have beliefs about the probability distribution of the utilities for each actions in a choice set, hence can be calculated for any choice set over which the rational optimization assumption can be applied.

The total amount of expected regret an individual feels at any point in time will be her *cumulative expected regret*, the expected regret attached to all past actions. Hence cumulative expected regret calculated at period  $t$  will be  $D(t) = \sum_{i=0}^t \delta^{t-i} d(i)$ , where  $0 < \delta \leq 1$  is some optional geometrical discounting factor. Finally, *coherence* will be the state of having cumulative expected regret equal to zero.

The key assumption of this model will be that individuals make mental adjustments to minimize their level of expected regret:

**Assumption 1:** At each period, individuals will adjust their preferences and/or beliefs in order to allow for minimal cumulative expected regret over their entire sequence of past and prospective actions and choice sets.

In other words, individuals will attempt to attain a state of coherence by locating a set of preferences and beliefs for which the optimal set of actions available provides them with the lowest amount of expected regret. A state of coherence, if feasible, allows an individual to believe with certainty that she can calculate the very best plan for maximizing a destiny in life that she has chosen. This in turn allows the individual to view her life in a teleological manner, with all her actions moving her closer to this ultimate destiny.

If we leave things at that, this kind of “meta-optimization” would present difficulties. There could be degenerate ways of achieving coherence, such as being indifferent over all outcomes or adopting beliefs totally at odds with the observed world. Moreover, the search space for minimizing expected regret would be infinitely large, with any conceivable utility function or set of beliefs in play as candidates for the individual’s sense of self. This in turn leads to a number of other assumptions placing constraints on beliefs.

**Assumption 2:** Individuals beliefs cannot exclude facts that derive from direct observation or logical inference.

This is in a sense a weaker form of the “information assumption” found in conventional rational choice theories, whereby all beliefs must derive from direct observation or logical inference. If this limit did not exist, individuals could achieve coherence by living in fantasy worlds where expected regret-inducing aspects of the environment simply disappear.

**Assumption 3:** Each individual's utility function will be constrained so to have a linear relationship of fixed slope with personal material welfare.

This assumption normalizes utility functions so that individuals cannot eliminate expected regret completely by adopting a "yogic" utility function in which all out-

comes have equal utility. Such a utility function, needless to say, would make predicting action impossible. The assumption furthermore ensures that expected regret-reduction cannot occur simply by dividing the utility function by an arbitrarily large number.

**Assumption 4:** The parametric form of individual utility functions and beliefs are limited to the sum or conjunction of those forms that are available from ideologies found in the individual's cultural environment. However, any coefficients in an ideology can be shifted or negated arbitrarily by and individual in the process of preference and belief adjustment..

This may need a bit of more explanation. Basically, what it is positing is that no variable or functional form of relationship between variables can appear in a belief or utility function unless there is some explicit ideology that is the individual has been exposed to that states that this variable or relationship between variables is significant.

For instance, if an individual is to have a utility function  $U = \dots + x + \dots$ , then there must be an ideology that states that the good  $x$  is a useful thing to have. However, once the ideology makes  $x$  available to the utility function of an individual, the individual can choose to disvalue the good, i.e.  $-x$  and/or place any weight upon it she wishes, i.e.  $2x$ .

**Assumption 5:** The process of minimizing expected regret will involve considering first the variables and variable functions found within ideologies in a sequence based on the salience of the ideologies in the cultural environment, where salience is a matter of the frequency with which an individual is exposed to communications espousing that ideology.

What this means is that those ideological messages that are broadcast more often at an individual, either through the mass media or through personal contacts, will have precedence in shaping that individual's preferences and beliefs. This does not mean, of course, that individuals will passively absorb such ideologies. An ideology may be rejected outright (i.e. not incorporated in preferences and beliefs) if its incorporation does not lead to reduction in expected regret. Likewise, even if it is incorporated, manipulation of coefficients may lead to an opposite affective disposition towards an outcome or belief in an opposite causal relation than that espoused by the ideology.

This model is not at all meant to be a purely formal construction, but rather to adopt and integrate concepts and findings from a number of different literatures in multiple social science disciplines. Quite briefly, the formal definition of regret used here in versions of regret theory that have been proposed in economics (Savage 1972; Loomes and Sugden 1982; 1987; Bell 1982), the main difference being that the earlier versions have used regret as a model of decision-making, rather than preference and belief formation. Indeed, in its substantive implications, the model is closer to psychological theories of dissonance and their close cognates (Aronson 1992; Harmon-Jones 1999; Cooper 2002). While dissonance has been defined in varying and contradictory

ways, newer versions of it in particular focus on the importance of action that seems to go in some way against an individual's own priorities or values. Likewise, the model is in line with recent sociological thinking on the construction of identity, particularly literature which focuses on the role of narrative and memory in creating a sense of collective purpose out of the past (Ricoeur 1988; 1992; Zerubavel 2004; Olick 2003; 2007). A number of other ties can be made, including narrative psychology (McAdams et al. 2006) and symbolic interactionism (Denzin 2007).

### 3 Implications for Preference and Belief Change

To begin, the focus will be on preference change that consists of increasing or decreasing weights on variables in linear utility functions. In other words, given a utility function in the form  $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ , where  $x$ 's represent variables and  $\alpha$ 's represent their weights, the focus will be on how shifts in  $\alpha$ 's affect expected regret.

**Proposition 1:** An individual can reduce expected regret for a past or intended action by raising the utility coefficient of a variable that is believed to be positively linked causally to the action.

Let the variable  $x$  stand for a good or outcome which, by prior belief, is thought to be causally linked to the chosen action  $a_1$ . Without loss of generality, we can assume  $x=1$  if the action is chosen and  $x=0$  if not (this simply represents a choice of measurement units). The most natural interpretation of  $x$  would be the intrinsic value placed upon the action itself or its direct consequences. In this case, increasing the weighting of  $x$  by some amount  $\Delta\alpha$  will shift the distribution for  $U_{ij} - U_{ij}$  to the right by  $\Delta\alpha$  units for all actions  $i \neq 1$ , in which case  $d^* = -\int_s \max(U_j^* - U + \Delta\alpha, 0) \phi_s(s) ds$ . It is straightforward from then on to show that this will reduce expected regret compared to  $d^* = -\int_s \max(U_j^* - U, 0) \phi_s(s) ds$  if there is at least some state of nature with positive probability density such that  $U_j^* - U > 0$ .

This theorem conforms with psychological findings on commitment, which reveal that individuals tend to increase their relative liking for alternatives after they have decided upon them (Brehm 1956). Commitment phenomena are often applied to explain the effectiveness of such influence techniques as "low-balling," where individuals are provided incentives to make choices, have some of their original incentives unexpectedly withdrawn, yet stick to their original choices because of post-decision increases in their intrinsic valuation (Cialdini, 1986). Finally, the theorem can be used to explain the "functional autonomy of motives," whereby means chosen to achieve particular ends are eventually valued intrinsically.

It also sheds light on the contradictory mechanisms Elster calls "sour grapes" and "forbidden fruit is sweet" (Elster 1983, chap. III). The first occurs when individuals adjust their preferences to adapt to their perceived circumstances (i.e. to increase their

expected utility), while the second occurs when adjustment is in the opposite direction. While Elster does not attempt to specify when these each of these opposing mechanisms will occur (and indeed believes it is impossible to do so), the theorem provides some indication that the effect will depend on the extent to which an individual has chosen to challenge the status quo in pursuit of new set of circumstances. Acceptance of the status quo breeds sour grapes for those new circumstances, while an act or intention to challenge will generate forbidden fruit.

As we move to beliefs, we can define the *net provision* of a good for an action as the amount of the good provided by that action compared to the highest amount provided by all actions available in the corresponding choice set. In this section, the focus will be on belief change that involves changing the expected value and the perceived variation of in the net provision of a particular variable for a chosen action.

**Proposition 2:** An individual can reduce expected regret for an action by increasing the expected net provision of a variable with positive utility weighting.

Let  $y$  stand for some good or outcome with positive utility weighting  $\alpha > 0$ . Suppose the individual in question changes her belief set so that the distribution of  $\hat{y}_{ij}$  shifts over  $\Delta \hat{y}_i$  to the right for all states of nature  $j$ . Then, assuming the individual has no other beliefs which force her to simultaneously change the perceived distribution of other variables, the perceived distribution for  $U_{ij} - U_{ij}$  for all actions  $i \neq 1$  is shifted  $\alpha \Delta \hat{y}_i$  to the right. The rest of the proof of theorem proceeds similarly to that for Theorem 1. Similarly, it can be shown that it is expected regret-reducing to reduce expectations of the net provision of a good with negative utility weighting.

This kind of belief adjustment corresponds to post-decision "bolstering", in which individuals adopt beliefs that emphasize the expected benefits that will the choice they have made, vis-a-vis the choice(s) that they have passed up (Knox and Inkster 1986; Younger, Walker, and Arrowood 1977). This can also be used to explain the "illusion of control" effect, where individuals have been found to systematically exaggerate the influence their choices have over important outcomes (Langer 1975).

These types of effects are relevant for the "wishful thinking" and "unwishful thinking" phenomena in which individuals adjust their beliefs either towards or against expectation of favorable outcomes (Elster 1983, chap. IV). The above theorem can specify the conditions under which each will occur. For an individual who has taken risks that depend on favorable circumstances occurring, expected regret can be reduced by wishful thinking, while for individuals who play it safe expected regret can be reduced by unwishful thinking.

## 4 Conclusion

This chapter has focused on simple processes of individual preference and belief change, but the coherence model can be used to explain both more complex processes of change at the individual level, such as the effect of exogenous structural changes on the extent of preference and belief formations, as well as collective processes of (for more detailed discussion, see Chai 2001, chap. 3). By applying these kinds of ideas it is possible to make predictions on such disparate outcomes such as the choice of economic (Chai 1998) and defense (Chai 1997a) ideology, as well as the formation of ethnic groups, their boundaries, and their propensity for violent collective action (Chai 1996; 2005). For ideology, the role of political activists, and particularly “ideological entrepreneurs” (Popkin 1979; Taylor 1989), is crucial in broadcasting messages that determine the alternatives that members of a population will consider first in altering their preference and beliefs in attempts to regain their coherence during times of rapid structural change. Those ideologues who propose preferences and belief changes that address the expected regret problems of the population will be most successful in getting their ideas adopted. For ethnic groups, the most important factors are the processes by which groups of individuals develop a collective identity and begin to mobilize on the basis of that identity. This is a multifaceted process that involves the acquisition of altruistic preferences (Collard 1978; Margolis 1982) and common social norms (Hechter and Opp 2005).

Overall, the coherence model is designed to provide an alternative to the conventional rational choice models that incorporates cultural, ideology, and identity into preferences and beliefs, but does so in a way that retains the generality, determinacy, and parsimony of the original model. Of course, there are bound to be multiple, and undoubtedly superior, models that can fulfill these criteria, but it is important for the long-term development of rational choice analysis that more attempts be made in this direction.

## References

1. Allport, Gordon W. (1937) The Functional Autonomy of Motives. *American Journal of Psychology* 50:141-56.
2. Andreoni, James and Miller, John (2002) Giving according to GARP: An Experimental Test of the Rationality of Altruism. *Econometrica* 70(2): 737-53.
3. Aronson, Elliot (1992) The Return of the Repressed: Dissonance Theory Makes a Comeback (Special Issue). *Psychological Inquiry* 3(4):303-11.
4. Bell, David E. (1982) Regret in Decision-making under Uncertainty. *Operations Research* 33:961-81.

5. Bolton, Gary E. and Ockenfels, Axel (2000) A theory of equity, reciprocity and competition. *American Economic Review* 100:166-93.
6. Brehm, Jack W. (1956) Post-decision Changes in Desirability of Alternatives. *Journal of Abnormal and Social Psychology* 52:348-89.
7. Chai, Sun-Ki (1996) A Theory of Ethnic Group Boundaries. *Nations and Nationalism* 2(2):281-307.
8. Chai, Sun-Ki (1997) Rational Choice and Culture: Clashing Perspectives or Complementary Modes of Analysis. In Richard Ellis and Michael Thompson, eds., *Culture Matters*. Boulder, CO: Westview Press.
9. Chai, Sun-Ki (1997a) Entrenching the Yoshida Defense Doctrine: Three Techniques for Institutionalization. *International Organization* 51(3):389-412.
10. Chai, Sun-Ki (1998) Endogenous Ideology Formation and Economic Policy in Former Colonies. *Economic Development and Cultural Change* 46(2):263-90.
11. Chai, Sun-Ki (2001) *Choosing an Identity: A General Model of Preference and Belief Formation*. Ann Arbor: University of Michigan Press.
12. Chai, Sun-Ki (2005) Predicting Ethnic Boundaries. *European Sociological Review* 21(4):375-391.
13. Charness, Gary and Matthew Rabin (2002) Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*, 117:817-869
14. Cialdini, Robert B. (1986) *Influence: Science and Practice*. Glenville, Ill: Scott, Foresman.
15. Collard, David (1978) *Altruism and Economy: A Study in Non-Selfish Economics*. New York: Oxford University Press.
16. Cooper, Joel (2007) *Cognitive Dissonance: 50 Years of a Classic Theory*. Newbury Park, CA: Sage Publications.
17. Cote, James E. and Charles G. Levine (2002) *Identity, Formation, Agency, and Culture: A Social Psychological Synthesis*. New York: Lawrence Erlbaum.
18. Denzin, Norman K. (2007) *Symbolic Interactionism and Cultural Studies: The Politics of Interpretation*. New York: Wiley-Blackwell.
19. Elster, John (1983) *Sour Grapes: Utilitarianism and the Genesis of Wants*. Cambridge: Cambridge University Press.
20. Fehr, Ernst, and Klaus M. Schmidt (1999) A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114(3): 817-68.
21. Harmon-Jones, Eddie (1999) *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*. New York: American Psychological Association.
22. Hechter, Michael and Karl-Dieter Opp, eds. (2005) *Social Norms*. New York: Russell Sage Foundation.
23. Knox R.E. and J.A. Inkster (1968) Postdecisional Dissonance at Post Time. *Journal of Personality and Social Psychology* 8:319-323.
24. Langer, Ellen J. (1975) The Illusion of Control. *Journal of Personality and Social Psychology* 32:311-28.
25. Loomes, Graham and Robert Sugden (1982) Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *Economic Journal* 92:805-24.
26. Loomes, Graham and Robert Sugden (1987) Some Implications of a More General Form of Regret Theory. *Journal of Economic Theory* 41(2):270-87.
27. Margolis, Howard (1982) *Selfishness, Altruism and Rationality: A Theory of Social Choice*. Cambridge: Cambridge University Press.
28. McAdams, Dan P., Ruthellen Josselson, and Amia Lieblich, eds. (2006) *Identity And Story: Creating Self in Narrative*. New York: American Psychological Association.
29. Olick, Jeffrey K. (2003) *States of Memory: Continuities, Conflicts, and Transformations in National Retrospection*. Durham, NC: Duke University Press.
30. Olick, Jeffrey K. (2007) *The Politics of Regret: On Collective Memory and Historical Responsibility*. New York: Routledge.

## The Coherence Model of Preference and Belief Formation

31. Popkin, Samuel (1979) *The Rational Peasant: The Political Economy of Rural Society in Vietnam*. Berkeley: University of California Press.
32. Rabin, Matthew (1993) Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83(5):1281-1302.
33. Ricoeur, Paul (1988) *Time and Narrative* (vols. 1-3). Chicago: University of Chicago Press.
34. Ricoeur, Paul (2006) *Memory, History, Forgetting*. Chicago: University Of Chicago Press/
35. Savage, Leonard S. (1972) *Foundations of Statistics*, 2nd revised ed. New York: Dover Publishing.
36. Taylor, Michael (1987) *The Possibility of Cooperation*. Cambridge: Cambridge University Press/
37. Younger, J.C., L. Walker and A.J. Arrowood (1977) Postdecision Dissonance at the Fair, *Personality and Social Psychology Bulletin* 3:284-287.
38. Zerubavel, Eviatar (2004) *Time Maps: Collective Memory and the Social Shape of the Past*. Chicago: University Of Chicago Press.