**Title**
Accommodation dynamics for comparing utilities with others



**Journal**
Philosophical Studies, 175(10)

**ISSN**
0031-8116

**Authors**
Narens, Louis
Skyrms, Brian

**Publication Date**
2018-10-01

**DOI**
10.1007/s11098-017-0966-6

Peer reviewed

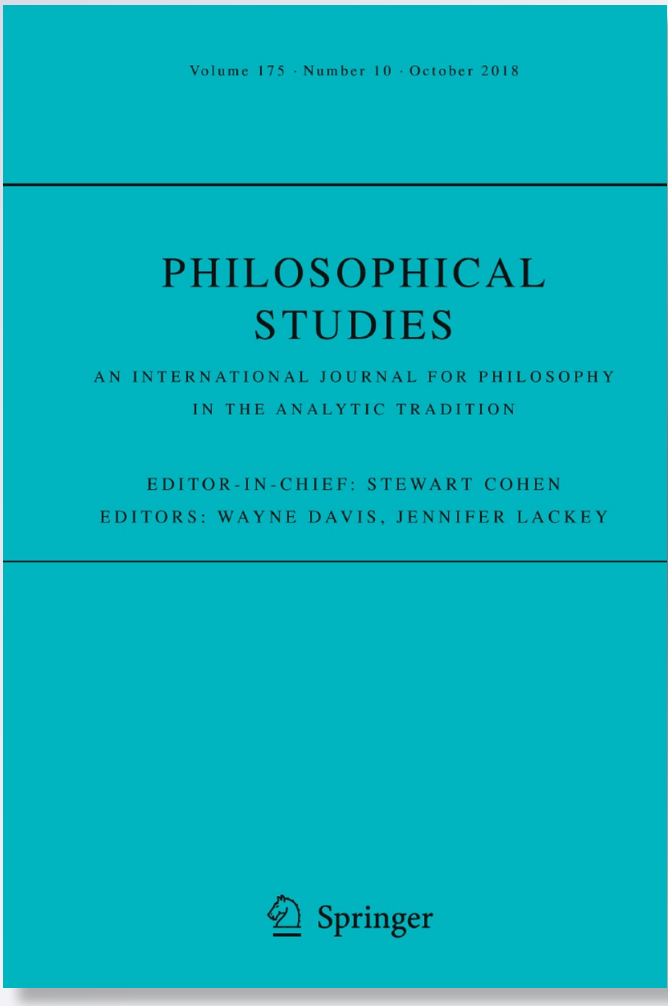
# *Accommodation dynamics for comparing utilities with others*

**Louis Narens & Brian Skyrms**

**Philosophical Studies**
An International Journal for Philosophy in the Analytic Tradition

ISSN 0031-8116
Volume 175
Number 10

Philos Stud (2018) 175:2419-2427
DOI 10.1007/s11098-017-0966-6



Philos Stud (2018) 175:2419–2427
https://doi.org/10.1007/s11098-017-0966-6

# Accommodation dynamics for comparing utilities with others

**Louis Narens**[1] · **Brian Skyrms**[2,3]

Published online: 14 August 2017


**Abstract** In interactive situations, agents can "learn" something that is not a preexisting truth. They can converge to an arbitrary convention, or tacit agreement. Once established they may even view it as an objective truth. Here we investigate accommodation dynamics for interpersonal comparisons of utility intervals. We show, for a large class of dynamics, convergence to a convention.

**Keywords** Interpersonal comparisons · Utility · Learning dynamics · Convention

## 1 Introduction

How do we compare our utilities with those of others? Early Utilitarians worried about the question, as have contemporary philosophers. The modern theory of utility, widely used in statistics and economics, provides no basis for this. In fact, measurement arguments involving scale invariance (e.g., Robbins 1935, 1938) have led most economists to believe that such comparisons have no place in economics. Nevertheless, we seem to make such comparisons all the time. Adam volunteers to wash the dishes even though Eve would do so otherwise, because he judges that the difference for him is smaller than the difference for Eve, who really hates doing it.

✉ Brian Skyrms
bskyrms@uci.edu

Louis Narens
lnarens@uci.edu

1 Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92717, USA

2 Department of Logic and the Philosophy of Science, University of California, Irvine, Irvine, CA 92717, USA

3 Department of Philosophy, Stanford University, Stanford, CA 94305, USA

He may just do it for Eve's sake. Or he may anticipate a future *quid pro quo*. We focus here on such comparison of differences. How is this to be explained?

Narens and Luce (2008) offered a skeptical hypothesis, consistent with both expected utility theory and ordinary practice. Adam and Eve, through long association, have come to an interpersonal equilibrium. There is no real underlying truth about correct interpersonal comparisons supporting this equilibrium. Many equilibria are possible. But Adam and Eve *behave as if* there were an underlying truth, and indeed may believe that there is. To put it another way, interpersonal comparisons of utility rest on an evolved *convention* between Adam and Eve. Such conventions are also possible for groups of more than two. They may be important for teamwork, collective action, or group agency. But how do we get to such an equilibrium? This is a question of dynamics, which is the subject of this paper. The closest work to the present one is Binmore's treatment of "empathy equilibria" (Binmore 1994, 1998), which bases interpersonal comparisons on evolved conventions involving social indices. The setting, however, is somewhat different. We do not use social indices. And Binmore's individuals can renegotiate the social contract at will behind a veil of ignorance.

Narens and Luce already raise the question of dynamics that can lead to such equilibria, and take some initial steps. Here we look at such accommodation dynamics in more detail, we examine cases in which they do not converge, and we characterize a broad class of such dynamics that provably converge to an equilibrium that supports interpersonal comparisons.

## 2 Adam's and Eve's utilities

There is a set of outcomes that Adam and Eve both care about. Each has coherent preferences over probabilities over outcomes in the standard way. (At this point we assume egoistic preferences. That is to say they are just about dish washing, and do not yet incorporate the other-regarding elements that we are trying to make intelligible.) This gives them each orthodox von Neumann–Morgenstern utilities such that preference goes by expected utility. Each of them has a utility scale family that is like a temperature scale family in that their scales have an arbitrary 0 value and an arbitrary unit. Knowledge of the scales by themselves does not allow Adam and Eve to make the common sense judgment with which we started. 0 points are not a problem, because they want to compare differences: is Adam's decrement in utility due to his washing small in comparison to Eve's increment in utility due to not washing. Comparing differences, choice of 0's for the scales wash out. Units are the problem. Choose different units and Eve's difference can be arbitrarily greater or smaller than Adam's. We can assume perfect knowledge of each other's preferences—they know each other well—but this does not help: Each knows each other's utility scale up to 0 and unit, but there is nothing to know about the other's units so they don't know how to make the trade-off between their utility differences.

## 3 Harsanyi's Gambit

But they *do* make judgments about the tradeoff. Harsanyi (1953) showed that this can fit into the framework of orthodox expected utility theory by supposing that they have extended preferences. Adam now has coherent preferences over an expanded outcome set, consisting of old outcomes being Adam's and old outcomes being Eve's. He then has an expected utility representation over all these outcomes on the *same scale*. This provides Adam with means to tradeoff of his units with Eve's. Similarly, Eve has extended preferences over the same outcome set providing her with the means to tradeoff her units with those of Adam. The problem is now that these comprehensive utility scales for these two actors may not agree on tradeoffs.

We could suppose that we are all the same under the skin, so that at some ultimate level of description Adam and Eve will agree on tradeoffs. Harsanyi says this. It is treated as a postulate. This is, to say the least, debatable. We do not want to assume this. We will assume, however, that Adam and Eve can talk to each other about their extended preferences, and their different views of tradeoffs.

## 4 Narens and Luce

Adam and Eve can talk to each other. They can observe miscoordinations. They can discuss them: "But I thought that you really, really minded washing the dishes." "No. not really. I don't enjoy it, but it isn't so bad." Adam and Eve may come to a *modus vivendi* as a result of these interactions. They may come to interact as if they could reliably make interpersonal comparisons of utility. They may come to believe that that can make true interpersonal comparisons. But such a view is mistaken, because other *modi vivendi*—other equilibria in interpersonal comparison—could have been achieved. People who hold this view have been misled into believing the objective validity of interpersonal comparisons of utility, even though such comparisons have no objective validity. Narens and Luce raise the question of equilibration dynamics for interpersonal comparisons in a preliminary way, and observe that different starting points may lead to different equilibria. They ask:

(i) how to characterize a broad class of plausible dynamic rules and
(ii) how to characterize the class of experiences that ultimately lead to equilibria. (p. 257)

## 5 Accommodation dynamics

Here we pursue this investigation of dynamics.
Adam might judge that it may not make much of a difference to Eve but a lot to him. Eve may contrarywise judge that it may not make much of a difference to Adam, but a lot to her. Then there is no coordination of who does the dishes. Even though each wants to respect the other's utilities, they disagree on the tradeoffs. The disagreement might go in the opposite direction, as illustrated in O' Henry's story

"The Gift of the Magi." Adam and Eve may both want to do the dishes because each thinks that the other cares more than they do.

It will be useful to regiment our description of the problem. Adam and Eve only care about comparing differences of their utilities with their perceived utilities of the other. Adam has a scale $u$ that is an expected utility representation over all the extended outcomes and Eve has a similar scale $v$ over her extended outcomes (=Adam's extended outcomes). These are only unique up to the choices of 0 and unit, but this doesn't matter. It is assumed that Adam has outcomes with maximum and minimum utilities, denoted, respectively, by $A_{\max}$ and $A_{\min}$, and Eve has similar outcomes denoted by $E_{\max}$ and $E_{\min}$. $u$ defines Adam's view of the comparison of his normalized units and Eve's. Specifically,

$$a_0 = \frac{u(E_{\max}) - u(E_{\min})}{u(A_{\max}) - u(A_{\min})} \tag{1}$$

defines a tradeoff constant $a_0$. This number is a ratio of differences, so we get the same number whatever scale of Adam's we use—that is, if instead of $u$ in Eq. 1 another scale $ru - s$ ($r > 0$ and $s$ is an arbitrary real) is taken from Adam's scale family, then Eq. 1 will still hold. We described this by saying $a_0$ *is absolute for Adam*. Likewise, Eve's extended utility function gives her an absolute tradeoff constant, $e_0$, for converting her units to Adam's. (Note that this is the reciprocal of the constant that she would use for converting Adam's utilities to hers.) So we have alternative views on how to convert Eve's utilities to Adam's.The distance between Adam's and Eve's tradeoff constants, $\delta_0 = |a_0 - e_0|$ is thus also absolute.

When Adam and Eve disagree about tradeoffs, they might accommodate by changing their extended utility functions so as to move their tradeoff constants closer together. If Adam moves close to Eve, and Eve moves closer to Adam, we say that they both evince *good will*. They might just split the difference or each might each take some different weighted average that give the other some positive weight. On the other hand, both may lack good will, and move further away from the other. This might, for instance, be for strategic reasons. Each wants to exploit good will of the other. In this case, the outcome will not be a tacit agreement, bit more likely a divorce in the Garden of Eden. Or perhaps Eve evinces good will, while Adam behaves strategically. The we may get a either a divorce or a tacit agreement, depending on just how Adam and Eve move. We do not what to just think of averages with fixed weights. The moves might be slow or fast. They might fluctuate in speed, with first relative intransigence and late accommodation. We would like to have a treatment that gives a general sufficient condition for convergence to an equilibrium of an accommodation dynamics that covers all these cases. For this, we just assume that if there is any difference, they will, depending on the difference, move closer together to reduce the difference. This is formulated formally as follows: For the first disagreement Adam moves $a_0$ in the direction of $e_0$ (but not equaling or surpassing it) by choosing a positive real number $p$ that does not depend on Adam's scale and making $a_1 = pa_0$ his new tradeoff constant. Because $a_0$ is absolute and $p$ does not depend on Adam's scale, $a_1$ is absolute. Similarly Eve obtains an absolute tradeoff constant $e_1 = qe_0$. Resolving other disagreements then lead to the sequence of absolute tradeoff constants, $a_i, e_i$, as $i$ ranges over the nonnegative integers. This gives rise to an accommodation dynamics that maps the difference $|a_i - e_i|$ into the smaller difference $|a_{i+1} - e_{i+1}|$ except when $|a_i - e_i| = 0$.

Formally, call a function $f$ from the non-negative real numbers into the nonnegative real numbers an *accommodation function* if and only if

(i) $f$ is continuous,
(ii) $f(d) < d$ if $d > 0$, and
(iii) $f(0) = 0$.

An *accommodation dynamics* has an accommodation function $f$ and starts from an initial positive real $d_0$ producing the *accommodation sequence*, $d_0$, $d_1 = f(d_0)$, $d_2 = f(f(d_0)), \ldots$.

We assume that $\delta_i = |a_i - e_i|$ is an accommodation sequence with accommodation function $g$ such that $g(\delta_i) = \delta_{i+1}$. Theorem 1 shows that this sequence converges to an equilibrium in which there is agreement on interpersonal comparisons of utility, that is, converges to $\delta = 0$.

**Theorem 1** *The sequence* $\delta_i$ *converges to an equilibrium in which* $\delta = 0$.

*Proof* The sequence $\delta_i$ is an accommodation dynamics. By (ii) and (iii) $g$ has 0 as its only fixed point, i.e., 0 is the only solution to $g(x) = x$. Because it is decreasing and bounded by 0, it has a nonnegative limit $r$. Because $\lim \delta_{i+i} = \lim \delta_i$, $r$ is also the limit of the sequence $\delta_{i+1}$. But $g(\delta_i) = \delta_{i+1}$. Thus $r$ is the limit of the sequence $g(\delta_i)$. But, because $g$ is continuous, $\lim g(\delta_i) = g(r)$. Thus $r = g(r)$, making $r$ a fixed point of $g$. Because 0 is the only fixed point of $g$, $r = 0$. □

For an averaging example of an accommodation dynamic, suppose that for each $i$, at the $i$th stage Adam moves $\frac{1}{8}$ of $|a_i - e_i|$ in Eve's direction and Eve moves $\frac{2}{8}$ of $|a_i - e_i|$ in Adam's direction, and $g(x) = \frac{5}{8}x$ is the accommodation function, but as we have emphasized, this is just a special case.

For each $i$, $(a_i, e_i)$ is a proper subinterval of $(a_0, e_0)$. Thus by Theorem 1, the $a_i$ and $e_i$ converge respectively to a points $a$ and $e$ in the interval $(a_0, e_0)$, where $a = e$. Because the $a_i$ and $e_i$ are absolute, $a$ (=$e$) is absolute. By appropriately choosing real $p_i$ and $q_i$ so that $a_{i+1} = p_i a_i$ and $e_{i+1} = q_i e_i$, convergence to any point in $(a_0, e_0)$ can be achieved.

The accommodation dynamics and Theorem 1 given above can be generalized in a number of ways. We could have used monotonicity rather than continuity for a technical requirement that covers a slightly different set of dynamics. We could have set the distance at which accommodation stops at some small positive number rather than zero, because it may be close enough for Adam and Eve. The argument is basically the same.

## 6 Three's company

Suppose that Adam and Eve have agreed that Eve's utilities are multiplied by 2 to convert to Adam's. Now they are joined in the garden by Susanne. Adam interacts with Susanne and they agree that Adam's utilities are multiplied by 2 to convert to Susanne's. Consistency then requires that Eve's utilities should be multiplied by 4 to convert to Susanne's. But Eve and Susanne, interacting separately, may have come to agree on a different tradeoff constant, say 1 to 3. If they all talk together, the inconsistency becomes apparent.

They may restore consistency in various ways. Suppose that the tradeoff between Adam and Eve is already fixed by habit. Then Adam and Eve are on the same scale. The question to be answered is where to put Susanne on that scale. If Susanne's utilities are multiplied by $k$ to convert to Adam's, they will be multiplied by $2k$ to convert to Eve's. If they all talk, they all know this. They are now in quite a different situation than previously, where they interacted separately.

Adam, Eve and Susanne may have different opinions about where Susanne should be put on this scale. These opinions can be expressed in terms of conversion to Adam's utilities, as 3 tradeoff constants, $k_1$, $k_2$, $k_3$. They can now accommodate by repeated weighted averaging—each taking a repeated weighted average that gives everyone's opinions some weight.

If instead, they all moved into the garden at once, with arbitrary extended preferences, their accommodation problem would be more complex. Now for Eve's tradeoff with Susanne, there are 6 numbers in play, the tradeoff from Susanne's point of view, the tradeoff from Eve's point of view, and the composite tradeoffs through Adam from the 4 combinations of points of view. You can visualize these numbers on the $X$-axis, the 6 numbers for Eve–Adam on the $Y$-axis, and those for Adam–Susanne on the $Z$-axis. There is a minimum and a maximum on each axis, and these define a box. The distances between minimum and maximum on the axes, $d_1$, $d_2$, $d_3$, give the dimensions of this box. The box contains all the tradeoffs. Perfect accommodation will shrink each dimension to zero.

Call the distance here the maximum of $d_1$, $d_2$, $d_3$. Then the argument using the accommodation dynamics works here as in the case of 2 players. It also works just as well for arbitrary finite numbers of players as for 3. If at every step the players responsible for the extreme values modify their views to move these values in the direction of the average, they will converge to an equilibrium.

This line of thinking leads to the following conclusions:

1. A general class of accommodation dynamics converges to a dynamic equilibrium in the case of $n$ agents. as before, there are different possibilities for who accommodates and how.
2. But if $n$ players have already formed a consistent set of tradeoffs, and one new player enters the scene, then the problem is simpler. The existing players already fit on one utility scale, and the question is just where to place the newcomer. Then simple weighted averaging serves as an example of accommodation dynamics.

## 7 Self-deception and accommodation dynamics

Narens and Luce suggest that comparability of utilities is based on self-deception. Individuals evolve conventional trade-offs, and then mistake them for objective fact. After all, the propensity of humans to mistake their own conventions for objective reality is well-known to anthropologists. Accommodation dynamics can provide some insight on the evolution of such conventions. Narens and Luce propose the following empirical study for the self-deception: Find a three's company situation where each pair has reach an equilibrium that they believe is objective, but, when taken together, the three equilibria reveal that the situation is impossible if the equilibria are objective. This is done by showing that if all pairs are objective then the extended utility data from any two pairs analytically predicts the data from the third pair. Narens and Luce reason that in situations where the three participants never simultaneously interact, the evolution trade-offs of two pairs will not specify the evolution of the third. This leaves the possibility—which Narens and Luce hypothesize as likely—that the remaining pair will not evolve in the manner that was analytically specified.

The problem with this proposal is finding such data. Our accommodation dynamics provides a rich framework that suggests experimental paradigms that likely would produce this kind of data as well as suggesting other experimental questions about the dynamics that individuals actually use.

We have seen that accommodation dynamics can be straightforward for pairs of individuals who use interpersonal judgments to tradeoff utility differences, and who know each other well enough to communicate their tradeoff constants. Belief in interpersonal comparability may lead to conventions of interpersonal comparison. Accommodation can be seen as a means for correcting for mistakes. Alternatively, individuals may not be deceived and may view themselves as negotiating a convention rather than discovering a truth.

Arriving at a convention becomes more difficult, but not impossible, as the group becomes larger. It is easier for a group that slowly adds members. In large groups formed all at once the difficulty may not not be in finding a suitable accommodation dynamics, but rather in everyone knowing each other's tradeoff constants. Such difficulties need not, in themselves, undermine the individuals' belief in interpersonal comparability. They may not notice inconsistencies in large groups. If they do, they may reason that there is not enough group interaction to correct for mistakes, or that some members of the large group may be falsifying their reports, or both. They may be satisfied with a coarse-grained approximate convention. Accommodation does not depend on the Narens–Luce self-deception hypothesis being true, but in the light of what we have found regarding accommodation dynamics, the self-deception hypothesis is still a viable possibility.

## 8 Conclusion

Whether peoples' utility can be validly compared is a central issue in philosophy and economics. In modern economics, the dominant view is that they cannot be compared. The intuition for this and related views was nicely summarized by Jevons (1887) in 1871:

> The reader will find, again, that is never, in any single instance, an attempt made to compare the amount of feeling in one mind with that in another. I see no means by which such comparison can be accomplished. The susceptibility of one mind may, for what we know, be a thousand times greater than that of another. But, provided that the susceptibility was different in a like ratio in all directions, we should never be able to discover the difference. Every mind is thus inscrutable to every other mind, and no common denominator of feeling seems to be possible. (p. 21)

The usual argument today against the intercomparisions assumes utilities come about through von-Neumann–Morgenstern expected utility theory and thus are represented on an interval scale. It is then argued that because there is no principled way of identifying units across different people's utility scales, valid intercomparisons are impossible (e.g., Robbins 1935, 1938). The usual argument for intercomparibility is that examples of it are observed all the time (e.g., Little 1957). This is the view taken by many moral philosophers. We find both kinds of arguments are deficient: The "against" doesn't explain why the rationality or the psychology inherent in von Neumann–Morgenstern cannot be extended in ways resulting in valid interpersonal comparisons. The "for" argument fails to consider alternative explanations for observed agreement behaviors regarding utilities that do not rely on an underlying truth.

Here we show that even given the best environment for interpersonal comparisons using extended utilities, a wide class of accommodation dynamics can account for observed behavior. They converge to a dynamic equilibrium at which individuals agree, but there are an infinite number of such equilibria and which one is reached depends on the starting point and the dynamics. The equilibria are invariant under change of scale. They provide a principled way of identifying scale units and constants throughout the accommodation process as they converge to an invariant interpersonal equilibrium.

This equilibrium can be viewed as a social contract and such contracts can increase the overall utility for the individuals involved by encouraging collective action (e.g., see Skyrms 1996). We believe that it is more productive to view interpersonal comparisons as social contracts instead of as a means for trying to achieve some social optimization or to carry out some moral imperative. Having many alternative equilibria is not an impediment for achieving an agreeable social contract.

**Acknowledgements** Research for this article was supported by Grant SMA-1416907 from National Science Foundation. We would like to thank Kent Johnson, Cailin O'Connor and Kevin Zollman for helpful comments.

## References

Binmore, K. (1994). *Game theory and the social contract I: Playing fair*. Cambridge, MA: MIT Press.

Binmore, K. (1998). *Game theory and the social contract II: Just playing*. Cambridge, MA: MIT Press.

Harsanyi, J. (1953). Cardinal utility in welfare economics and the theory of risk taking. *Journal of Political Economy*, *61*, 434–435.

Little, I. M. D. (1957). *A critique of welfare economics* (2nd ed.). London: Oxford University Press.
Jevons, W. S. (1887). *The theory of political economy*. London: MacMillan and Co.
Narens, L., & Luce, R. D. (2008). Why we may have been mislead into believing in the interpersonal comparability of utility. *Theory and Decision*, *15*, 247–260.
Robbins, L. (1935). *An essay on the nature of significance of economic science* (2nd ed.). London: MacMillan.
Robbins, L. (1938). Interpersonal comparisons of utility. *Economic Journal*, *48*, 635–641.
Skyrms, B. (1996). *Evolution of the social contract*. Cambridge: Cambridge University Press.