

Norms and Rationality. Is Moral Behavior a Form of Rational Action?

Prof. Dr. Karl-Dieter Opp
Universität Leipzig (Emeritus)
University of Washington, Seattle (Affiliate Professor)
opp@sozio.uni-leipzig.de



I. The Problem

The question to be addressed is

Can rational choice theory (RCT) explain the effects of norms?

The question is often formulated in the following way:

Is moral behavior (=norm following) „rational“?

Problem 1: „Rational“ has numerous meanings!

Problem 2: Even if „rational“ is clear: a simple **classification of behaviors as „rational“ is not interesting**. Important is the question whether RCT can **explain** norm following behavior.



There are two answers to this question.

The **incentives thesis**: norms are incentives (i.e. costs and benefits) for action and, thus variables of RCT.

The **autonomy thesis**: norms are distinctive from costs and benefits and can thus **not** be integrated into RCT.

Here are some quotations that illustrate the theses:



Proponents of the incentives proposition:

“ ... internalization of a norm will mean that an individual will come to have an internal sanctioning system which provides punishment when he carries out an action proscribed by the norm or fails to carry out an action proscribed by the norm“
(**Coleman** 1990: 293).

In the work on the **dictator and ultimatum games** fairness norms decisions of the subjects are explained by RCT, and fairness norms are a major explanatory variable.



Proponents of the autonomy thesis:

“.. duty, or obligation, has no place in an economic, or utilitarian, explanation. Things we do because of duty, or obligation, are things done irrespective of benefit or outcome ...”

(**Udéhn** 1996: 87-88)

Etzioni (1986: 162) argues that "normal people do some things because they are right, whether or not they enjoy them."

March and Olsen (2006, see also 1989) distinguish between the "logic of appropriateness" and the "logic of consequentiality." They adopt a "vision of actors following internalized prescriptions of what is socially defined as normal, true, right or good, without, or in spite of, calculation of consequences and expected utility"

Esser (2010 KZfSS, special issue), **Elster** (Cement: 125, 127)



Before the two theses are discussed it is useful to specify the definition of norms because this concept is used in different meanings.

How the concept of norms is used in this

paper:
(1) A **norm** is defined here as any statement asserting that something ought or ought not be the case under certain conditions.

(2) This paper refers to norm **internalization**: A norm is **internalized** to the extent that following the norm is an intrinsic motivation.



II. Contents of the Presentation

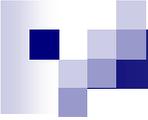
It may be the case that the validity of the incentives or autonomy thesis depends on the **version of RCT** that is applied.

It is therefore useful at the beginning to set out the **version** of RCT that will be used in this paper

AND

to analyze its **implications** for the incentives and autonomy theses.

This then is the basis for the discussion of the **arguments in the literature** in the second part of the paper/presentation: the question will be what the version of RCT is that the authors use.

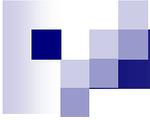


Part 1: Norm Following and the Wide Version of Rational Choice Theory

- What Rational Choice Theory is All About
- Norm Internalization as Preference: A Defense of the Incentives Thesis

Part 2: The Arguments for the Autonomy Thesis (PAPER)

- The Irrationality of Norm Following: Jon **Elster**'s Arguments – **IN PRESENTATION**
- Rational Choice Theory's Assumption of Mono-Utility (“Pleasure”): Amitai **Etzioni**'s Case for a Multiple-Utility Conception – **IN PRESENTATION**
- The Logic of Appropriateness and the Logic of Consequentiality: J.G. **March** and J.P. **Olsen**'s Approach
- **Moral Action vs. Utility Maximization**: The Fatal Blow to Rational Choice Theory?



III. Norm Following and the Wide Version of Rational Choice Theory

What RCT Is All About: The Wide Version

- **Preferences** (or, equivalently, desires or goals) are conditions for behavior.
- **Constraints or behavioral opportunities**, i.e. events that allow or prevent achievement of the goals, determine behavior.
- **Utility maximization**: Actors choose the behavior that they think is best for them, i.e. is the best way to achieve the goals.

Example: giving a tip in a restaurant; goal or preference: appear as a generous person; constraint: availability of money for tip.



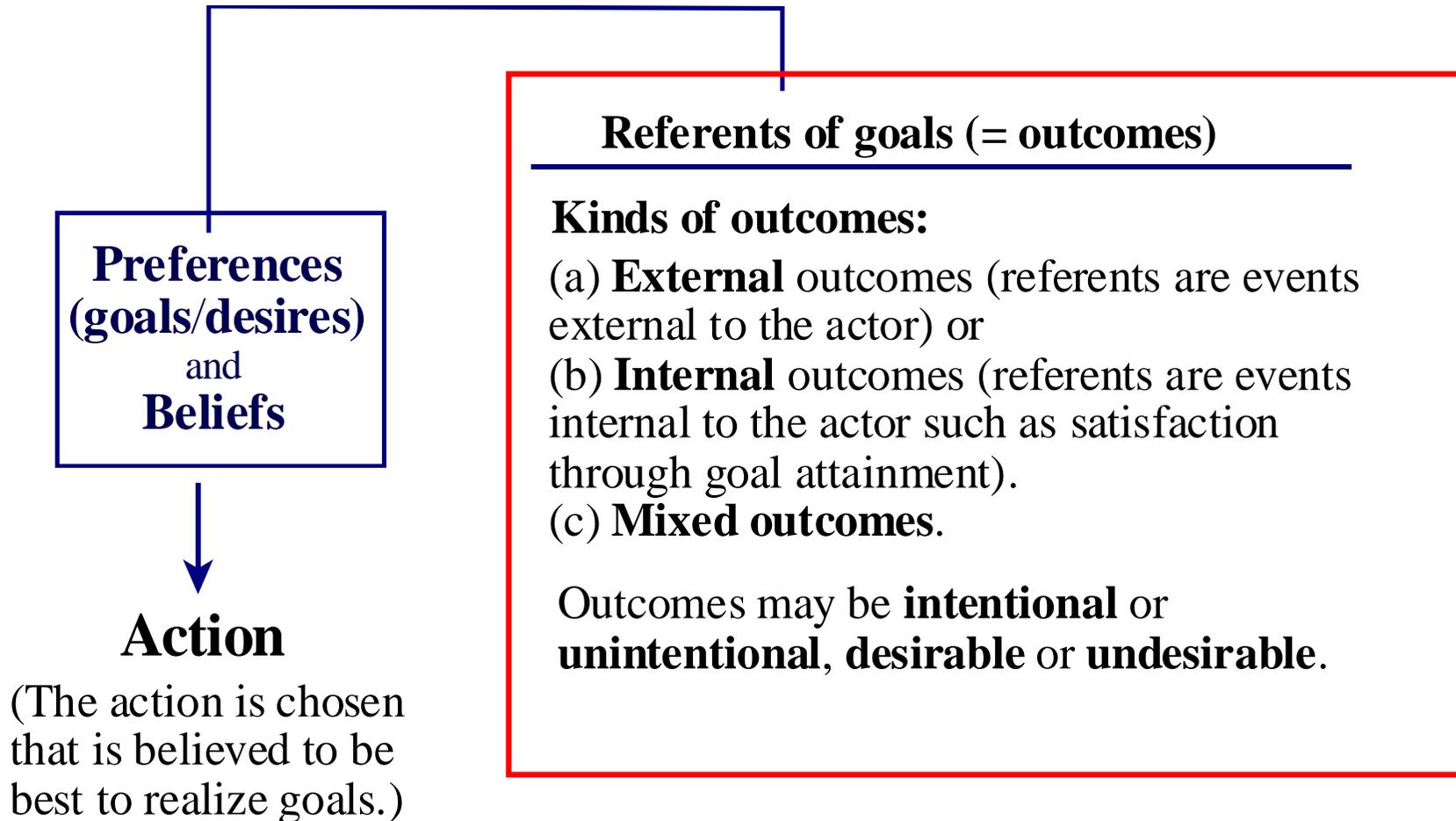
A wide version of RCT is used:

- **All kinds of preferences** may determine behavior (i.e. altruistic as well as egoistic motivations are possible).
- **Perceived constraints** and not objective constraints are relevant = **beliefs**.
- There are **no restrictions on the kinds of beliefs** (they may be wrong, incomplete etc.).
- **No calculation** is assumed: the claim is that action is governed by preferences and constraints. Thus, habits or spontaneous behavior are consistent with the wide version.

Important implication (relevant for the following discussion):

Preferences (or goals/desires) refer to “**outcomes or end states**” (Aarts and Dijksterhuis 2000: 54). These are the **referents of the goals**. In other words, if an action is performed and the goal is achieved, this state is then the “**outcome**,” i.e. the causal effect of the action. (**Example:** action “giving a tip”; goal “appear as a generous person”) – this is the “outcome,” i.e. the effect of the action.)

There can be different kinds of outcomes:





What does it mean that a norm – such the felt obligation to give tips – is internalized?

It means that there is a **goal** (or preference or desire) to follow a norm (i.e. actors „seek to fulfill obligations ...“, March and Olsen 2006: 689).

Question: Are there „**outcomes**“ if internalized norms are obeyed or broken?

Following norms brings about different kinds of **outcomes**:
internal (having a good conscience etc., i.e. internal rewards)
external (giving a tip makes the waiter believe ...)

Breaking norms brings about different kinds of **outcomes** as well:
internal (bad conscienc etc., i.e. internal punishments)
external (stealing makes the victim worse off).



Conclusion

According to the wide version of RCT, the internalization of norms is a variable of RCT. This is in line with the [incentives thesis](#).

But this is controversial.

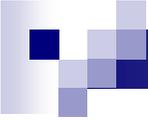
What follows are major arguments that are provided in the literature [for the autonomy thesis](#).



IV. The Irrationality of Norm Following: Jon Elster's Arguments

The argument (1989a: 99):

“**Rational action** is concerned with **outcomes**. Rationality says: If you want to achieve Y, do X. By contrast, I define **social norms** by the feature that they are not outcome-oriented. The simplest social norms are of the type: Do X, or: Don't do X ... Rationality is essentially **conditional** and **future-oriented**. Social norms are either unconditional or, if conditional, are not future-oriented.”



An analysis of the argument:

- Are norms **not outcome-oriented**?
 - What do norms as **statements** refer to? Simple vs. complicated statements (lying...). (Problem for RCT?)
 - Is obeying/breaking of **internalized** norms not outcome-oriented? (Does Elster refer only to external outcomes?)
- Is norm following **unconditional**? I.e. no behavioral alternatives are considered (spontaneous behavior).
 - To be sure, this holds for many (not all!) norms.
 - But this is not a problem for a wide version of RCT
- Is norm following **not future-oriented**?
 - Meaning of, e.g., „give tips in restaurants“ refers to future situations (norms hold in the past as well as in the future)!
 - Sometimes future outcomes are relevant – see lying example above.

Conclusion: Elster's arguments are not tenable.



V. The Assumption of Mono-Utility: Amitai Etzioni's Case for a Multiple-Utility Conception

KDO3 RCT is concerned with **pleasure (or satisfaction)** „derived from doing what one likes, enjoys, and finds pleasurable, distinct from

KDO4 the sense of affirmation that accompanies living up to one's **moral commitments** ... that are often in themselves taxing

KDO5 rather than pleasurable” (1986: 160 – emphasis not in the original).

Thus, RCT addresses **pleasure** but cannot adequately deal with **moral action**.

KDO3 **Genuß, Vergnügen, Lust**
Karl-Dieter Opp; 19.10.2010

KDO4 **Affirmation=Bestätigung**
Karl-Dieter Opp; 19.10.2010

KDO5 **taxing = anstrengend, schwierig**
Karl-Dieter Opp; 19.10.2010



Etzioni grants that pleasure and morality refer to **utility** which encompasses pleasure and morality.

However, if RCT would use an **overarching concept of utility** that includes pleasure and commitment, RCT would face serious **methodological problems**.

Thus: either one restricts RCT to pleasure or one incurs methodological problems by using a wide concept of utility.



Does RCT include only pleasure (like enjoying ice cream) and not utility in a wide sense?

No! RCT claims that people maximize **utility** in the sense to **enhance their well-being**. This may be fun (like eating ice cream) or performing a moral act.

Etzioni's **example** is: A mother "who dashes into to the fire to save *her* child" does not feel pleasure.

But: the mother is certainly better off after rescuing her child.



To support his claim that RCT deals with pleasure Etzioni cites **Jeremy Bentham**:

"nature has placed mankind under the governance of two sovereign masters: *pain* and *pleasure*" (Bentham 1982[1789]: 11).

This is, according to Etzioni, what RCT is about.

However, Bentham actually refers to utility in a wide sense:

utility is "that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered" (Bentham 1982[1789]: 12).



Etzioni's arguments: Why does a wide concept of utility violate basic requirements of concept formation?

1. A broad concept „ceases one's ability to explain.“

In general, thus, a concept should have „**explanatory power**“ or should be **theoretically fruitful** – i.e. make a theory true if included in a theory.

Employing only „pleasure“ would make the theory wrong in situations where moral motivations determine behavior. Example is voting or saving money (both are mentioned by Etzioni!).



2. With a broad concept of utility it is no longer possible to differentiate between the various kinds of motivations that affect behavior.

It is true, that the *theory* does not differentiate. But the *researcher* may classify motivations as he or she likes. Such differentiations are even necessary in **applications** of RCT is applied to explain behavior. For example, in explaining saving behavior (Etzioni's example) the factors that seem relevant must be measured. If donations are to be explained it is necessary to differentiate between "self-love" and love for others (163).



3. A broad concept of utility makes the theory "unproductive and tautological" because "whatever one does is said to 'reveal' one's pleasure."

Etzioni does actually refer to **circularity**. This means that evidence for the existence of the independent variable is „inferred“ from the existence of the dependent variable. (Popper example: why does it thunder? ...)

In a wide version – and in any theory – it is an obvious requirement that independent variables must be measured independently of dependent variables.

RCT is also clearly not **tautological** ...



4. How can behavior be explained if it is determined by multiple utilities?

Are the actors able to make a decision in situations where they face different kinds of utility?

The answer is clearly "yes": actors are able to **form "summary feelings"** (172): they apply a common measuring rod for the different kinds of utilities and engage in a comparison of the "calibrated" utilities.

For **example**, saving behavior is influenced by the size of the income, by the interest rates, by the desire to dispose of more money after retirement, and by a norm that it is a right thing to save (174-175).



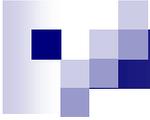
5. What motivates moral behavior – if not utility?

A question that is not answered by Etzioni is **what the motivation for norm-following** is if this does not make individuals better off.

Why does a mother dash into a fire to save her child – to take the example by Etzioni – if that is not a behavior that makes the mother better off or happier than if she would let the child die?

6. Conclusion

The existence of multi-dimensional utility is not a problem for RCT.

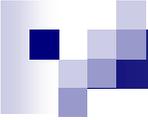


VI. The Logic of Appropriateness and the Logic of Consequentiality: March and Olsen's Approach

Two „logics“ are distinguished (March and Olsen 2006):

The **logic of appropriateness**. "Action ... is seen as driven by rules of appropriate or exemplary behavior ... Rules are followed because they are seen as natural, rightful, expected, and legitimate. Actors seek to fulfill the obligations encapsulated in a role, an identity ... processes of reasoning are not primarily connected to the anticipation of future consequences as they are in most contemporary conceptions of rationality" (690).

The **logic of consequentiality** assumes "self-interested and rationally calculating actors, instrumentalism and consequentialism" (691).



Do these types of “logics” exist in reality? Yes.

Is the „logic of appropriateness“ inconsistent with RCT?

It seems that this is not claimed by the authors:

- It is hypothesized that rules **solve problems of groups** by coordinating activities in a way “that makes them mutually consistent and reduces uncertainty” (695).
- The „**logics**“ **are not regarded as mutually exclusive**: In their book (1989: 25) the authors "could imagine political actors treating alternative rules and interpretations as alternatives in a rational choice problem." This suggests that RCT could adequately deal with both logics.
- An **alternative theory to RCT is not proposed**.

The **conclusion** thus is that there are situations where the two "logics" obtain. But this does not require different theories.



VII. Moral Action vs. Utility Maximization: The Fatal Blow to Rational Choice Theory?

Most advocates of the autonomy thesis claim that norm following is not utility maximizing.

Here is again the quotation from Udéhn 1996: 87-88:

“.. duty, or obligation, has no place in an economic, or utilitarian, explanation. Things we do because of duty, or obligation, are things done irrespective of benefit or outcome ...”

To what extent is this argument acceptable?



Problem 1: If norm following is not utility maximizing, what, then motivates people to obey norms?

There is no clear answer in the literature.

Often adherents of the autonomy thesis (1) **implicitly** assume utility maximization or (2) provide **vague hypotheses**.

Re 1: Elster (1989b: 106) addresses the "contrast between norms and self-interest" and states that "both types of motivation may enter into a single action ... When the norms require me to do X and self-interest tells me to do Y, I may end up with a compromise." Is not a "compromise" some behavior that is best for an actor?

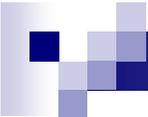
Re 2: Elster (1989a: 105) claims that internalized norms work like an "internal gyroscope that keeps them adhering steadily to norms, independently of the current reactions of others"
– hardly an illuminating statement.

Does all this not suggest that there is utility maximization?



Problem 2: If norm-related behavior is not utility maximizing, there should be no costs or benefits if norms are obeyed or broken. If there are such costs and benefits it is implausible that actors ignore them in their decisions.

1. There are **always** internal emotional costs and benefits (good/bad conscience, shame, satisfaction) when obeying or breaking norms.
2. Costs and benefits occur **in certain types of situations**.
E.g., if norms are generally accepted (e.g. not smoking at certain places) and if deviations are visible, costly external sanctioning is likely.
3. If there are **conflicts of normative and non-normative goals** (exceed speed limits or miss an appointment) a compromise is often found – this means that costs and benefits are considered.



Problem 3: Norms are used as rationalizations and, thus, norm-related behavior is not utility-maximizing.

Example: A wants B to stop smoking and appeals to the norm not to hurt others. A does not subscribe to the norm – it is used as a rationalization to bring about a certain behavior.

Does RCT have a problem with explaining why A invokes the norm?

Does RCT have a problem of explaining how B reacts?

The issue is whether obeying or breaking **internalized** norms is utility maximizing. But if norms are only used as justifications or rationalizations, there is no internalization of a norm!

Thus, utility maximization due to norm internalization is not predicted by RCT.

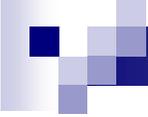


Problem 4: There are many examples indicating that norm compliance is against one's own interest and is thus not utility maximizing.

Examples: Vendetta/revenge (Elster 1990); rescue of Jews during WWII (Monroe 2004).

For these examples the possibility of applying RCT has been defended (for vendetta: Hamlin 1991; for rescuing Jews: Opp 1997, Varese and Yaish 2000 – see also Kroneberg, Yaish and Stocké 2010).

Novels and operas: For example, in Peter Ilyich Tchaikovsky's Eugen Onegin Lenski feels offended by his friend Onegin and demands "satisfaction." The friends sing a duet when they meet for a duel which includes the text: ".. we silently and coldbloodedly prepare to destroy each other. Oh, should we not burst out laughing before we stain our hands with blood, and should we not part friends? No! No! No! No!" Also: Don Carlo by Verdi.



Question: Do these examples provide evidence that norm-related behavior is not optimizing?

These examples seem to contradict optimizing behavior because the costs of the behaviors are tremendous and are visible for observers; the **benefits are not clear**.

However, a closer look at the situations may reveal that it is **not implausible** that actors do what is best for them from their perspective – see the defense of the incentives theses for the examples of vendettas and rescuing Jews.

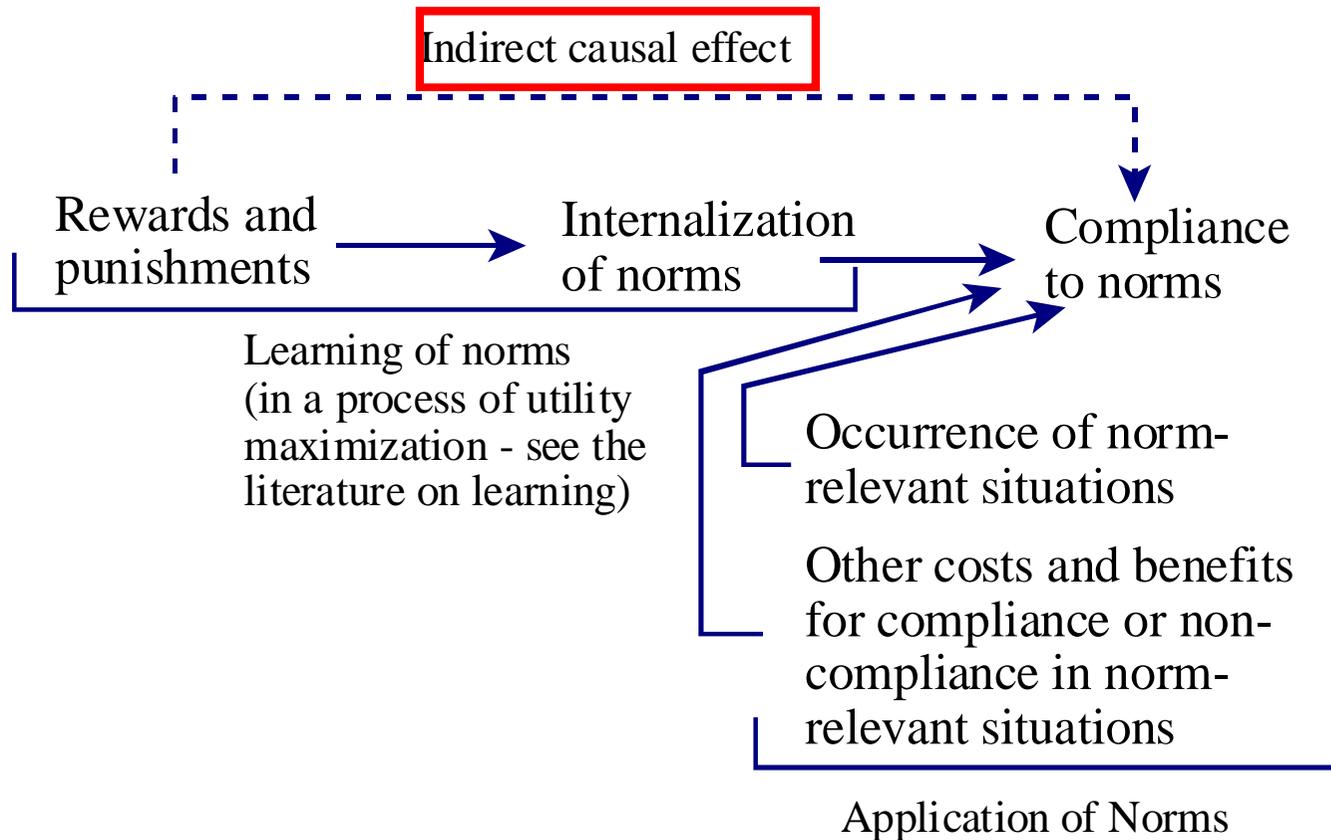
Anyway, it is not an argument against the incentives theses if one only mentions cases where the costs for a behavior are apparently extremely high. A **detailed empirical analysis is necessary**. The work defending the optimization thesis for vendettas and rescuing Jews suggests that there are also benefits of the respective behavior.



Problem 5: Learning norms happens by rewarding or punishing behavior. The norm is acquired in a process of utility maximization. Is it plausible that these rewards are no longer effective when it comes to obeying or breaking given norms?

Implication of the learning model: **prior incentives have an indirect effect on present behavior** – see the following figure:

Figure 2: The Process of Norm Internalization and Norm Compliance



Implication: if there is **strong internalization**, „Other costs and benefits“ are less effective. Thus: past learning (= optimization) affects later norm-relevant behavior.



Problem 6: Is the spontaneity of norm following an argument against utility maximization?

Many norms – especially in everyday life – are followed spontaneously: if the situation where the norm holds occurs the norm is enacted without calculation.

Examples: norms of politeness (greeting friends), table manners, behavior in restaurants (ordering meals, giving tips etc.) etc.

Many authors think that in these situations individuals do not maximize utility.

As was said before, RCT does not assume that people calculate, but this may be wrong.

Question: is there any evidence that spontaneous norm following is utility maximizing?



Spontaneous action and norm following originate in a process in which:

- **A first step is calculation**

- example: P moves to a new apartment and wishes to find the fastest way to his or her workplace.

- **A decision is made, i.e. a behavioral program is adopted.**

- P decides, e.g., to take the subway.

- **The behavior is enacted** to check whether the expected costs and benefits occur.

- Perhaps several options are tried out such as going by car.

- **If the calculation was correct the behavior will be performed without calculation.**

- Why? Calculation is costly!

- **If the cost-benefit ratio for the chosen behavior changes** the decision is re-considered and calculation occurs again.



Dual-process theories confirm this process.

See Chaiken and Trope 1999. More recent discussions and summaries are provided by Bargh et al. 2001; Osman 2004; Smith and DeCoster 2000; Strack and Deutsch 2004.

Conclusion: Existing evidence indicates that there is „rational spontaneity“: spontaneous norm following maximizes utility.



Problem 7: Do norms make groups better off?

So far the question is: is norm **compliance with given internalized norms** utility maximizing?

Even if that is granted the question is: is the **emergence of norms in a group** utility maximizing, i.e. do norms make **groups** better off?

The vast **literature on norm emergence agrees** that some factors are involved in norm emergence in a group:

- externalities;
- sanctioning;
- social networks.



Brennan and Buchanan (1985: IX) have summarized the thrust of the insights of social scientists in norm formation:

"At the most fundamental level, rules find their reason in the never ending desire of people to live together in peace and harmony, without the continuing Hobbesian war of each against all."

Conclusion: The literature on the emergence of norms suggests that in general norms emerge in a society because they increase individual well-being from the perspective of the actors involved

Problem 8: Could there be a direct test of the autonomy thesis?

Perhaps a factorial survey?



VIII. Conclusion

Two theses were discussed:

- **incentives thesis**

- norm following is utility maximizing

- **autonomy thesis**

- norm following is about doing one's duty and is not maximizing utility.

The discussion of various arguments in favor of the **autonomy thesis** indicates that it is **not tenable**.

- 
- The discussion is based on a **wide version of RCT**. In a narrow version in which only external outcomes are considered, only norms with external outcomes can be integrated into RCT. Why is RCT restricted to external outcomes? This question has never been answered by proponents of the autonomy thesis.
 - In general, authors who advance the autonomy thesis do not suggest a clear **alternative theory**.
 - Those arguments for the autonomy thesis were selected that seemed most plausible. So perhaps there are **other arguments** that have not been discussed in this paper and show that norm following is not utility maximizing?

This is – unfortunately (???) – the

END