

Karl-Dieter Opp*

What a Theory of Social Norms and Institutions Should Look Like

Experimental Economics, Rational Choice Sociology, and the Explanation of Normative Phenomena

<https://doi.org/10.1515/auk-2020-0013>

Abstract: In the previous issue of *Analyse & Kritik* (2020, vol. 42, issue 1) Alexander Vostroknutov (3-39) aims at a ‘synthesis’ of economics with ‘psychology, sociology, and evolutionary human biology.’ This paper argues that his approach needs to be complemented at least by work from sociologists and social psychologists. Starting with problems of defining and measuring norms it is then claimed that a theory of norms should address the origin, change and effects of norms and model micro-macro processes. This should also be the goal of a theory of institutions (which are defined here as sets of norms—norms in the sense of accepting oughtness statements). We show how the social psychological value expectancy theory can be applied to model the variety of incentives that could play a role in explaining the effects of norms. Regarding the origin Coleman’s theory of norms is applied to show how Vostroknutov’s dissatisfaction-norms hypothesis can be improved.

Keywords: externalities, experiments, institutions, mechanism explanations, measurement of norms, norms, value expectancy theory

1 Introduction

This article is mainly a comment on contributions to the focus on ‘experiments on social norms’ in A&K—2020 (1). We concentrate on Vostroknutov’s article (2020), but we also address general issues of a theory of social norms.

When explanations of normative phenomena are at issue the first question to be answered is what the meaning of the concepts of norms and institutions is because their usage differs in the literature. After discussing conceptual issues we deal with the measurement of norms. The remainder of this article is con-

*Corresponding author: Karl-Dieter Opp, Universität Leipzig, Leipzig, Germany (Emeritus), University of Washington, Seattle, USA (Affiliate Professor), e-mail: opp@sozio.uni-leipzig.de

cerned with explanatory questions. We start with the question of what a theory of norms should explain. We then turn to ‘approaches’ to explain normative phenomena—functional, evolutionary, and mechanism explanations. The next sections then discuss the effects and origin of norms.

2 How Should Norms and Institutions Be Defined?

There are numerous definitions of ‘norms’ and of the related term ‘institutions’ (for a summary see Opp 2015). In order to know what one is talking about it is necessary to specify how these concepts are used.

Two examples illustrate the different meanings of norms in the literature. (1) “A *norm* will be defined as an expectation about how one ought to act, enforced by the threat of sanctions or the promise of reward” (Kerr 1995, 33). (2) “... a social norm is (i) a behavioural regularity that is (ii) based on a socially shared belief how one ought to behave which triggers (III) the enforcement of the prescribed behaviours by informal social sanctions” Gächter/Fehr 1997, 276 – italics omitted).

These and many other definitions consist of the following definitional criteria. (1) Norms are defined as *expectations of others* how one ought to act. (2) Norms are defined as *behavioral regularities*. (3) A personal belief about how one *ought* to act is another definition of the norm concept. This refers to a perceived personal obligation (i.e. to the acceptance or internalization of an oughtness statement). (4) Often there are *complex definitions* consisting of at least two of those definitional criteria. Often *sanctioning* is included. In the first definition above a norm only exists if there is oughtness *and* (positive or negative) sanctioning (i.e. enforcement) in case of compliance or norm violation. The second definition above combines regularities *and* oughtness *and* enforcement. Definitions further differ in regard to the *object* of oughtness, expectations, and sanctioning. The previous definitions address *behavior*. But attitudes, goals and beliefs may be subject to norms as well. For example, Christians ‘must’ believe in the existence of god, and one ‘must not’ have negative attitudes toward women or Jews. For the sake of simplicity, in what follows we only deal with behavior.

Norms may be *micro variables* (individuals think they ought to behave in a certain way) or *macro variables* (there are shared expectations or laws).

It is striking that often authors who write about norms do not define the term and, if the term is defined arguments are missing why not another definition is chosen. The reader might explore the contributors to the A&K issue 1 (2020). Most of the time we do not find a clear definition of norms or a justification of why a certain definition is used. For example, Vostroknutov (2020, 4) addresses “social norms,

customs, conventions, moral rules, fashions, etc.” and “traditions.” Apparently, these terms have different meanings. Is it useful to deal with all these phenomena at the same time?

The question arises what the most useful definition is. It is preferable to define norms by only *one definitional criterion*. The reason is that a theory explaining a single phenomenon (such as behavioral regularity) refers to a broader class of phenomena than a theory that explains more than one phenomenon simultaneously (e.g. behavioral regularity *and* sanctioning in case of noncompliance). The latter class of phenomena (regularity *and* sanctioning) is a narrower class of objects than the former one (only regularity is the definitional criterion). A wider definition thus allows the explanation of a larger class of phenomena. Those definitional criteria that are excluded from the definition can be used as dependent or independent variables. For example, if norms are defined as beliefs that some behavior ought to be performed, sanctioning and behavioral regularities could be used as additional variables.

Another question is how to discuss an article if it is not clear what the definition is that the author applies. Even if no explicit definition is provided the context often suggests what definition is used. For example, Vostroknutov (2020, 25) seems to refer to the oughtness definition. This definition is explicitly introduced when he presents his own theory. In general, if there is no explicit definition in a paper we assume that the oughtness definition is used. That is to say, *norms are statements about what ought to be the case*. The dependent or independent variable thus is *acceptance of oughtness statements*. Norms in this sense are often the subject of the social sciences, ‘norms’ normally refer to ‘oughtness.’ Furthermore, it is interesting to explain the origin, stability, and effects of when people think a behavior is more or less obligatory.

The term *institutions* is also defined in different ways, as these two examples illustrate: “We define institutions as a system of human-made, nonphysical elements—norms, beliefs, organizations, and rules—exogenous to each individual whose behavior it influences that generates behavioral regularities.” (Greif/Laitin 2004, 635) “Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction” (North 1990, 3). According to the first definition, institutions refer to almost everything. In the second definition, institutions are defined as sets of rules. We adopt this definition: institutions are *sets of oughtness statements* such as a market order.

In what follows we concentrate on norms. The question whether explaining single norms or systems of norm requires a different procedure will not be discussed.

3 A Note on Measuring Norms

Vostroknutov is one of the economists who cares about measuring norms and also suggests difficulties (2020, 23-24, 29-30; see also Görge/Nosenzo 2020 in this issue of A&K). In sociology and social psychology, norms are normally measured in surveys by presenting respondents with statements to which they can more or less agree. For example, one can measure a norm to participate in protest action by items such as: “If there is a demonstration and I would not join, I had a bad conscience.” High agreement means high acceptance of a felt obligation to protest.

Norms are most of the time *conditional*: they hold only under certain conditions. For example, the norm is not ‘one must tell the truth.’ One must tell the truth only under certain conditions. There are, for example, white lies. The *factorial survey* (Auspurg/Hinz 2015; Jasso 2020; for an application see Jasso/Opp 1997) is most useful to measure the conditionality of norms. To illustrate, assume we want to know under which conditions a person should earn 3000 \$ (after taxes) per month. Assume we conducted some preliminary studies and found that conditions might be, among other things, the person’s education and age. We categorize both variables and construct descriptions of situations (called vignettes) that consist of all meaningful combinations of those values. For example, a vignette might describe a person with a BA in sociology (category of education) and with the age of 25 to 30 (category of age). The accepted norm is the rating dimension and could be a scale (e.g. ranging from -1 to +1, with zero in the middle), with the categories ‘not at all justified’ (-1) to ‘fully justified’ (+1). Each respondent is presented with several vignettes. This measurement technique is thus well suited to measure the conditionality of norms.

These procedures measure norms directly by asking people what their opinions are about what should or should not be the case. Often norms are measured indirectly (i.e. by proxies). That is to say, certain phenomena are measured that are supposed to correlate with the oughtness beliefs. For example, Kimbrough/Vostroknutov (2018a, 148, see also the discussion in Kliemt 2020, 74-76, and the discussion of ‘direct’ and ‘indirect’ measurement in Erkut 2020) inform experimental subjects that ‘the rule is to put the balls into the blue bucket.’ The authors use ‘the cost incurred following the rule as a measure of an individual’s propensity to follow norms in other environments.’ The propensity is thus not measured directly. The problem is that in case of a falsification of the tested hypotheses one might always argue that the proxy did not refer to the norm to be measured.

Is it necessary to measure norms at all? This is a strange question discussed at length by Görge/Nosenzo (2020). The question is odd because claims of a theory that certain factors influence other factors *must* be tested by empirical re-

search, and this is generally accepted by empirically oriented scientists. If, thus, norms—however they are defined—are among the explanatory variables it is obvious that their influence on the explanandum must be tested. If norms are the explanandum it is obvious as well that they *must* be measured. Sociologists in particular have developed numerous methods to conduct empirical research to test theoretical propositions.

4 Testing General Hypotheses with Laboratory Experiments

Findings referring to laboratory experiments with game theoretical designs (such as the Dictator and Ultimatum Games) raise the question to what extent these findings apply to natural situations. The existing findings for each experiment are *singular statements*. These are sentences referring to certain times and places. A summary of findings of all actually conducted experiments is a singular statement (or a set of singular statements) too. For example:

H1: For the Dictator Games experiments *that have so far been conducted* it holds: Proposers give more than 20% of their endowments to the responder.

However, it is often assumed that H1 holds for *all* dictator games, i.e. for all experiments that have been conducted in the past and that will be conducted in the future:

H2: For *all* Dictator Games experiments it holds: Proposers give more than 20% of their endowments to the responder.

Note that this statement goes beyond a summary of existing findings. H2 is a ‘generalization’ in the sense that a singular statement (H1) that refers to a *finite* set of a certain class of situations is assumed to hold for the respective *infinite* class of situations. It is important to note that the ‘generalization’ H2 is *not a logical inference* from the singular statement H1. Such ‘inductions’ are invalid, as is well known from formal logic. The transition from H1 to H2 is an *intuition* by the researcher. Thus, any generalization in the sense of a logical inference from the specific to the general statement is invalid.

Another question is whether H2 can be ‘generalized’ to *natural* situations. This is the question of the external validity of experimental findings, as it is sometimes formulated. However, there is no ‘*external validity*’ and, thus no valid ‘generaliza-

tion' to nonexperimental situations. That is to say, an inference (or 'generalization') from statements holding for some situations to statements referring to other situations is invalid (see particularly Gadenne 2011; Stroebe et al. 2018). To illustrate:

H3: Individuals give more than 20% of their yearly income to relatives who are in serious trouble.

This hypothesis does not follow logically from the previous hypotheses. What is possible, however, is that H2 *suggested the idea* or *led to the intuition* of a researcher that H3 *might* be valid. Such intuitions are perfectly legitimate. But no claim should be and can be made that there is any *logical* connection of H2 and H3 (or any hypothesis similar to H3). Whether the intuitions are valid must be determined by empirical research.

All this is well known to scholars with some background in formal logic and philosophy of science. But this group seems to be a minority. One of the indicators for this assertion is the frequent discussion about the possibility of 'generalizations' of experimental findings. The *procedure of finding propositions* that do not only refer to the specific experimental situations should be: (1) Start with general propositions before an experiment is conducted. In the present example, these hypotheses may refer to conditions that have an impact on donating some part of one's income to charity organizations or to people who are in need of financial support. (2) Explore logical implications of the general propositions: which results are to be expected in the experiments? (3) These hypotheses referring to expected experimental findings are then to be tested. The important point is to start with general propositions. For example, H3 could be tested with an Ultimatum Game in which the responder is more or less in trouble—he or she might be seriously ill, unemployed or may have lost his or her partner.

The frequent discussion of the 'generalizability' of experimental results indicates that researchers are not quite satisfied with experimental findings. This is understandable because, strictly speaking, experimental findings about, for example, Dictator or Ultimatum Games are completely irrelevant for natural situations. If they are used as heuristic devices for intuitions about more general hypotheses, the question arises: why are those hypotheses not formulated before the experiments are conducted?

Those social scientists who are interested in informative hypotheses that can be applied to a wide range of natural situations must conclude: the hypotheses tested in the experiments addressed in this focus of A&K are unsatisfactory. Their informative content is low. That is to say, the hypotheses can only be applied to a rather narrow class of (the experimental) situations, and they explain only a narrow class of phenomena (behavior in the specific experiments). What is needed is to get

inspired by the results of the already existing experiments and formulate general informative propositions, and then test them again. Before new experiments are conducted, the procedure described before should be applied.

5 What a Theory of Norms Should Explain

There are three theoretical questions that existing hypotheses about norms address. One is why norms come into being. A second question is about the effects of already existing norms. A third question is the change or stability of existing norms. This question can be answered by applying theories about the origin of norms: if the factors that lead to norms change, then norms should change as well. However, there are cases when those factors change but norms do not change. This is due to the costs of change. For example, a keyboard with a certain order of letters on the bars might have been invented to speed up typing. Due to the advent of computers there were no bars attached to letters anymore so that another distribution of letters on keyboards may allow quicker typing. But changes of keyboards in enterprises in an economy are costly so that a change of the keyboards may not be profitable (Lewin 2002). There is ‘path dependence’ in the sense that past behavior or institutions influence present behavior and institutions. In explaining the stability or change of norms there may thus be other factors than those which led to the origin of norms.

It would be ideal if a theory of norms describes *social processes* (i.e. mechanisms) that show the single steps of norm emergence and of the impact of norms in a group. Since norms are often macro properties *micro-macro modeling* is useful. For example, incentives that are brought about by institutions such as markets or laws by governments (macro properties) lead to individual incentives for a certain behavior (micro properties). These incentives change individual behavior (micro properties) which is then aggregated to a macro explanandum. An illustration is the so-called Coleman boat—which is illustrated by Figure 1 below (Coleman 1990a, chs 10 and 11).

When explaining the *effects of norms* one question is to what extent existing norms are *obeyed or violated*. Norms have further *side effects*. If norms come into being they often have unforeseen and undesirable effects. For example, laws that protect handicapped employees by impeding dismissal has the side effect (which is by most observers regarded as undesirable) that employers hire fewer people with a handicap because the law makes this costly. A theory of norms should be capable to explain those side effects. If those side effects are behaviors the economic model can explain them. Accordingly, the side effects of welfare measures are discussed in economics.

A theory of norms should have, as every theory, a high *informative content* (Popper 1959, ch. VI). Intuitively, a theory of norms should be capable of explaining a large number of very *specific* features. For example, it should be capable to explain when tipping customs change in what way and why these customs differ across countries (see, e.g., Azar 2020). A high informative content is further given if a theory can be applied to a large number of situations, i.e. if its *range of application* is relatively extensive. A theory of norms should thus explain the origin and effects of all kinds of norms. Otherwise its informative content is relatively low.

Vostroknutov (2020, 4) argues that “not all social norms, customs, or traditions are suitable for ‘utilitarian’ economic analysis.” For example, assume dressing up Christmas trees (the example is mentioned by Vostroknutov 2020, 4) is a norm. Compliance or non-compliance is an action and, thus, should be amenable to ‘utilitarian analysis.’ That is to say, the economic model should be capable of explaining compliance or non-compliance to any kind of norm. Otherwise, i.e. if compliance to certain kinds of norms can not be explained, there is a limited range of application of the theory and, thus, a relatively low informative content.

Many theoretical statements about social norms have a very low informative content. They are rather *orienting statements* (Merton 1957, 87–89). These are “broad postulates which indicate types of variables which are somehow to be taken into account rather than specifying determinate relationships between particular variables” (88; Homans 1967, 14–18). Take the hypothesis that norm following depends on selfish as well as non-selfish motives. It is not specified which kinds of motives lead to following norms and when violation of norms occurs in what way.

6 Functional, Evolutionary and Mechanism Explanations of the Emergence and Effects of Norms

There are different *theoretical approaches* to explain norms. These are not clear theories that specify exactly general relationships between variables. In the A&K focus functional and evolutionary approaches are insinuated in the article by Vostroknutov (2020, e.g. 5–6, 24–25). What do these approaches contribute to explain normative phenomena?

Functional explanations are common in economics and in other social sciences. In sociology they were advanced in particular by functionalists such as Talcott Parsons. The major idea was that norms originate when they contribute to the equilibrium or survival of a society (or group). Vostroknutov (2020, 5) advances a

similar proposition: “social norms ... emerged ... as devices that simplify ingroup cooperation and by doing that increase the survival chances for people who follow them.” This kind of theory has been criticized extensively and convincingly already a long time ago (see in particular Hempel 1959; Nagel 1956). One problem is what ‘survival’ or ‘equilibrium’ of a group mean. For example, does the absence of survival of a tribe mean that each member died? Is survival still given if 51% of the members leave the group? Explaining the origin of a social phenomenon by its effects is a very serious mistake. To use Vostroknutov’s (2020, 5) example: “washing your hands is a good way to prevent the spread of diseases. Given its benefits to the community, we can expect that such a norm can become common.” The procedure is to explain a phenomenon (washing hands) that does not yet exist by its effects that have not yet occurred. To be blunt, this has nothing to do with a sound scientific explanation.¹

It seems, however, that authors who use a functionalist argument actually mean that *actors perceive* the ‘functions’ as *behavioral consequences* that are incentives for their behavior that is performed after the incentives originated. Only this is an acceptable causal explanation that is compatible with the economic model (see the discussion of value expectancy theory below). Accordingly, Vostroknutov’s example can be reconstructed in the following way. (In a personal communication about a former version of this paper Vostroknutov suggested that the functionalist argument actually refers to the following explanatory argument that is perfectly compatible with an application of the economic model.) The following mechanism explains the origin of the norm of washing hands. At the beginning of the process no one washes hands. Then somebody starts this practice and notices that his or her diseases become less frequent or disappear. This person tells others about the benefits of hand washing. Because avoiding diseases is in general a high benefit, others imitate this behavior. As the practice of hand washing spreads, not following this practice becomes a negative externality (a cost to others), if the diseases spread by infecting others one has contact with. This externality could be a condition for the emergence of a norm to wash hands (see the discussion of Coleman’s theory below). The example illustrates one possible process of the emergence of a *spontaneous order*. It resembles the emergence of a nonsmoking norm (which is mentioned in Vostroknutov’s comment as well, see also below).

Note the difference between this explanation and a functionalist explanation. The former describes incentives that lead to a behavior, which, in turn, changes

¹ This is still not known in the literature. In a recent contribution Eriksson 2020 regards functional explanations as a legitimate form of explanations. But the basic critiques, e.g., by Hempel and Nagel, cited before, are not discussed.

incentives for others which again influences behavior etc. The functionalist explanation would argue that the effects of hand washing (which have not yet occurred) cause the behavior (hand washing) that has not yet occurred as well. The functionalist thus ignores the mechanism of the emergence of the norm.

Evolutionary explanations often assume that phenomena develop so that 'evolutionary success' or 'efficiency' or the good of the group originate. These explanations are as problematic as functional explanations. However, this does not hold for evolutionary work in the tradition of Boyd and Richerson (e.g. 2005). They are actually similar to mechanism explanations (see below).

Applying the economic model these arguments must be changed in the following way. Not the actual effects of norms matter, but the extent to which the existence of a norm that does not yet exist has certain *expected* (and perceived) beneficial effects for actors; and these perceived consequences set in motion processes that lead to the emergence of norms. These norms may but need not be efficient or make the group better off. Actors may make mistakes and their behavior may generate inefficient norms.

To illustrate, a *functionalist explanation* of the norm of hand washing would be that such a norm that does not exist would increase the health of a group if it exists. The norm therefore originates. An alternative explanation by *applying the economic model* is the following. Assume that at a certain time people do not wash their hands. Now let there be research that shows that hand washing diminishes various illnesses. Let people believe this and want to stay healthy. Thus, the overall utility of washing hands (compared to not washing hands) increases in a society. These individual actions then aggregate to a 'hand washing custom' (i.e. to regular behavior). A consequence is a greater health in the society. A functionalist might disregard the individual processes and simply see the 'function' of washing hands. Actually, the 'function' refers to previously expected behavioral consequences.

This kind of explanation is a *mechanism explanation* (Hedström/Swedberg 1998; Opp 2005; see also 2013b). Mechanism explanations refer to social processes that explain the origin, stability and effects of norms and institutions by applying the *wide economic model*. This assumes, very briefly, 'bounded rationality' (i.e. limited cognitive capabilities of actors) and subjective (and not objective) utility maximization; all kinds of preferences and subjective constraints (i.e. beliefs) are admitted. They must be determined by empirical research. Value expectancy theory to be discussed below illustrates the wide version of the economic model. The basic arguments for applying this wide version are discussed in Opp (1999, see also Opp 2020a).

One type of mechanism explanations are micro-macro explanations. The macro variables are social events (such as new research results) or institutions (such as private property rights). These macro variables affect individual incen-

tives for certain kinds of behavior. These behaviors then aggregate to some macro property. The hand washing example shows this. More complicated examples are the explanations of the failure of nations (Acemoglu/Robinson 2012) and of the rise of the Western World (North/Thomas 1973).

7 Modeling the Impact of Normative and Non-Normative Preferences on Norm Following

The contribution by Vostroknutov (2020) is a clear exposition of the experimental work of economists that focuses on material *and* ‘social preferences’ (i.e. nonmaterial preferences). He thus applies the wide rational choice model. The prime example in which both preferences matter for the decisions of the actors is the Dictator Game. They can be explained if a fairness norm is assumed: proposers who can distribute any amounts of a certain endowment to a responder usually do not keep all the money. These results suggest that subjects in experiments follow norms (or, in general, social preferences) at least to some extent.

Two questions arise. (1) What exactly are the preferences that determine behavior in the experiments *and in natural situations*? (2) If a general theory is applied: what are the propositions of this theory?

(1) The original distinction of preferences in the Dictator Game (see Vostroknutov 2020) was between material (i.e. monetary) preferences (i.e. a preference to get material goods including money) and normative preferences (preferences to heed norms such as a fairness norm). If we assume, according to the economic model, that the *kind* of norm has an influence on the *kind* of behavior one needs to determine which *kinds* of norms the actors accept. This follows from the economic approach (or, equivalently, rational choice theory): following an accepted norm provides a benefit (such as a good conscience), whereas violating an accepted norm is a cost (such as a bad conscience or shame). Explaining the behavior of the proposer in the Dictator Game thus requires knowing exactly what the behavior is the ‘fairness norm’ refers to. For example, the proposer could think that he should *share half* of the endowment with the responder. Another norm could be to give only *a fourth*—with the justification that he or she—the proposer—owns the endowment and, thus, it is justified to keep the major part. Both norms will lead to different behaviors. To explain the kind of behavior the kind of norm must be known.

Vostroknutov (2020, 15) discusses the kinds of nonmaterial preferences in the Dictator Game and in similar games. He mentions that dictators might feel entitled to keep the money. It may be important if the responder is someone from

the ingroup or outgroup. This could mean that the *fairness norm is conditional*: under certain conditions one feels more or less obliged to transfer a certain amount of the endowment. Other incentives are expected ‘kindness’ or approval by others, negative sanctions, or intrinsic rewards from following reference persons. All this is consistent with Vostroknutov’s theoretical argument, but the kinds of incentives need to be specified in more detail. This requires detailed hypotheses and measurement.

(2) The subjects’ behavior in the Dictator and Ultimatum games is influenced by preferences. However, the economic model consists of preferences *and constraints*. In the games the constraint is the experimental setup. For example, the proposer has only a certain amount of money available, and the person who receives the money cannot punish the responder in the Dictator Game. Given the constraints, only the preferences determine behavior. But in more complicated situations it is necessary to consider the influence of the constraints as well. For example, assume in the Ultimatum Game the proposer has a preference for reciprocity (i.e. getting approval from the responder). This preference may exist even if the responder is unknown and will never be met again. The *expectation* (or, equivalently, the *subjective probability*) of getting the approval of the responder is thus the (perceived) constraint. This variable may be of central importance and must be included in explanations of behavior.

Another requirement for modeling the effects of norms is to consider the *intensity of the preferences*. Assume the preference for money is very strong, but the acceptance of a norm for sharing half of the endowment is weak. The effects of this constellation of preferences are different from a constellation with a strong norm and a weak monetary preference. The norm-dependent utility function in Vostroknutov’s paper (2020) takes this into consideration by addressing monetary incentives and normative valences.

If the economic model is accepted a specific version that allows a detailed modeling of preferences and perceived constraints (i.e. subjective probabilities) is the social psychological version of SEU theory, namely *value expectancy theory* (VET).² We will briefly describe this version because it is rarely applied in explaining the experimental results the articles in A&K focus on. The theory assumes that a behavior that is included in the *perceived* behavioral alternatives is becoming the *performed* behavior. For the sake of simplicity, assume the proposer in an Ultimatum Game has 2 Dollars at his or her disposal and that the options given by the experimenter are: keep the entire amount or give 1 Dollar to the responder.

² See, in particular, Feather 1982, 1990; Wigfield et al. 2016. Subjective expected utility (SEU) theory is a version of VET that is also used in economics. See Stigler 1950a, 1950b.

These are the behavioral alternatives. VET assumes that the behavior is performed that has a relatively large number of consequences that are valued relatively positively and subjectively expected to a high extent. This overall utility of a behavioral alternative is denoted as its subjective expected utility (SEU). The consequences (and their values and expectancies) are, in terms of the economic model, *incentives* or *costs* and *benefits*. VET claims that *the action with the highest SEU is performed*. There is thus subjective utility maximization—given the preferences and beliefs of the actors.³ To illustrate VET with this example, we write two behavioral equations (see below), one for each behavioral alternative. There are two behavioral consequences: getting the money and follow the norm. The actors assign utilities and subjective probabilities to each consequence. Let the utilities U range from -1 to +1 and the subjective probabilities p from 0 to 1. We assume that the utility of the material incentive is relatively high (.8 for 2\$ and .4 for 1\$), whereas the utility of abiding by the norm is only .2. The proposer knows that the probability of getting the 2\$ if he or she does not share (p_{k2}) is zero, and the probability of abiding by the sharing norm (p_{ks}) is zero as well. The SEU of sharing is higher so that sharing will be chosen.

$$\begin{aligned}
 SEU(Keepall) &= p_{k2} * U(2\$) + p_{ks} * U(Normtoshare) \\
 &\quad p_{k2} = 0; U = .8; p_{ks} = 0; U = .2; SEU = 0 \\
 SEU(Share) &= p_{s1} * U(1\$) + p_{ss} * U(Normtoshare) \\
 &\quad p_{s1} = 1; U = .4; p_{ss} = 1; U = .2; SEU = .6
 \end{aligned}$$

Note the following features of VET.

- (1) The theory is useful because it forces the scientists to ascertain in detail perceived consequences, utilities and subjective probabilities. These are usually not determined in detail in the games (they are all simultaneously contained in the payoffs and not separated as in VET) – see in particular Vostroknutov (2020).
- (2) No strategic behavior is necessary. The proposer might spontaneously give half of the endowment—maybe he or she always shares in certain situations his or her assets, and this situation is given in the game. So there is no calculation. It is an interesting question to determine to what extent actors in the game

³ Kliemt (2020, 79-80) argues that utility maximization, as in VET, is “classical, e.g. ‘Benthamite’ utility theory,” whereas modern utility theory holds that alternatives are ranked higher ‘for what reasons ever.’ We disagree with this claim. For a discussion see Opp 2020b.

act spontaneously or weigh explicitly the costs and benefits of the chosen behavior.

- (3) VET is a version of the wide rational choice theory. The latter assumes, as was mentioned before, subjective (and not objective) utility maximization, and admits all kinds of motives (including altruism and the goal to follow norms). In VET the preferences are the utilities, and the subjective probabilities are the perceived constraints. Subjective utility maximization in VET is the hypothesis that the perceived behavioral alternative with the highest SEU will be chosen. VET specifies the effects of the independent variables in more detail than the usual version of rational choice theory (for details see Opp 2020b).
- (4) The kinds of perceived behavioral options, the perceived consequences, their utilities, and subjective probabilities must be measured empirically. This is obvious because in an empirical science allegations need to be tested. This concurs with Vostroknutov's (2020, 18, 22) views who describes a detailed measurement procedure in one of the experiments he reviews. In the previous example, the values of the p 's and U 's might be quite different. So the question is how they are distributed in a situation. This must be determined by applying the methods of empirical research. VET does not provide hypotheses that predict the specific p 's and U 's. This is plausible because VET is a general theory, whereas the specific behavioral alternatives, utilities and subjective probabilities are singular statements (they are, in terms of the logic of explanation, the initial conditions of VET).
- (5) Let us look in more detail at the role of accepted norms in VET. A psychological fact is that violating an accepted norm is costly: it is linked to a bad conscience or shame. This is a consequence, in terms of VET, of the behavioral alternative 'rule following.' VET thus implies that there is only one kind of methodological individualism that includes internalized norms as one possible kind of incentive. The extent to which this matters needs to be determined empirically. This concurs with the position of Kliemt (2020): there is only one kind of methodological individualism that includes normative as well as non-normative incentives.

So far, we included only one norm in a behavioral equation. Norms are often related to other norms that are used as justifications. For example, the norm not to annoy others by smoking may be implied by a more general norm not to impose costs on others. A behavioral consequence of violating a nonsmoking norm may thus also generate a bad conscience by deviating from the general accepted norm. This would then be an additional cost of norm violation.

- (6) It is sometimes denied that norm following can be explained with rational choice theory. It is held that norms have nothing to do with 'rational' behavior, and only this is behavior that can be explained with rational choice theory. The latter refers to future-oriented behavioral consequences. Norms, however, are not future oriented: one does things because they are right and not because they have certain consequences. This is the position advanced in particular by Elster (1989; 1991). This argument is mistaken (for a detailed discussion see Opp 2013a). VET and, consequently, the wide economic model include norm following or violation as behavioral consequences. People want to avoid the violation of internalized norms because this *brings about* (i.e. leads in the future to) a bad conscience or shame.
- (7) Numerous other incentives (behavioral consequences with the respective probabilities and utilities) can be added in the equations. If, for example, behavioral options are 'compliance' and 'noncompliance' to a norm, behavioral consequences might be: punishment (i.e. its utility and subjective probability)—the punishment may come from authorities such as the police as well as from friends; compliance might lead to intrinsic rewards from imitating others or from a norm of imitation (children should do what their parents do); there might be a general norm that is complied to that justifies the specific norm; there may be numerous non-normative preferences and, thus behavioral consequences (violating a speed limit might save time). Kliemt (2020, 70) mentions that performing an action may be of intrinsic value. There is thus a consequence that a behavior fulfills an autonomous need to do something. VET would model the intensities of the need and the likelihood that the respective action fulfills the need.
- (8) VET is a micro-theory that can be embedded in a *micro-macro model*. This means that incentives are influenced by macro phenomena. For example, institutions of a market order are macro variables. They specify private property rights that allow to realize various goals such as to set up enterprises. These individual actions aggregate to macro behavior such as the supply of goods.

The structure of such a micro-macro model that explains the effects of norms is summarized in Figure 1. Norms and institutions are macro variables. They are related to some kind of collective action which is a macro variable as well. On the macro level there is only an indirect causal effect that is generated by a process that links the macro and micro level. Norms or institutions (such as private property rights) lead to incentives for members of a group that lead to a certain behavior (such as setting up an enterprise). These individual behaviors aggregate to some macro property such as a certain supply of goods. Note that there are pre-existing incentives. Among them are internalized personal norms as micro properties. There

are other goals (such as getting a higher income) that are not influenced by the macro properties. Figure 1 is based on the so-called Coleman boat (Coleman 1990a, chapters 10 and 11).

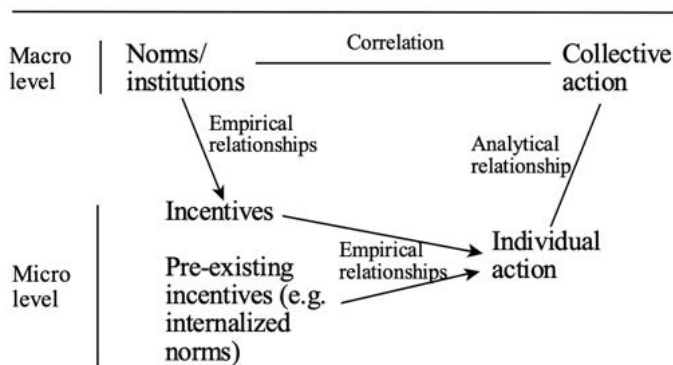


Fig. 1: The Basic Model for Explaining the Effects of Norms and Institutions

8 The Origin of Social Norms

Vostroknutov (2020, 25-34) proposes a theory that refers to the literature of experimental economics.⁴ One of the major theories in sociology is a theory about norm emergence by James S. Coleman (1990a, chs 10 and 11, see also 1990b). Although the theory is not without problems (for a recent discussion with further references see Opp 2018a), its major hypotheses seem to be a valid explanation of the emergence of social norms. In what follows this theory is first briefly exposed. In a next step Vostroknutov's theory is described and compared with Coleman's theory.

⁴ He summarizes an unpublished paper by Erik O. Kimbrough and himself (A Theory of Injunctive Norms, mimeo, Chapman University and Maastricht University). The paper can be downloaded. I compared this paper with the summary in the A&K issue and think this is a good summary of the paper.

8.1 Coleman's Theory of Social Norms

Social norms, as Coleman defines them, “specify what actions are regarded by a set of persons as proper or correct, or what actions are improper or incorrect” (1990b, 37 – the following references refer to this article). He thus uses the oughtness definition. ‘Sanctioning’ refers to actions of the norm beneficiary (e.g. a nonsmoker) aimed at influencing the target’s (e.g. the smoker’s) actions.

Coleman’s explanation proceeds in two steps. The first is the *demand for a norm*. This refers to a regulatory interest, i.e. the wish that a norm comes into being. The second step is the *realization of a norm*. This means that the norm is generally accepted and backed by sanctions.

The trigger of norm emergence are *externalities*. These are behaviors of actors that are costly or beneficial to other actors. Pollution or being exposed to smokers are examples for negative externalities. A house owner who cleans the sidewalk provides a positive externality to passers-by. Those who are exposed to negative externalities are interested in reducing them; those who are exposed to positive externalities want them to last, but those who provide those externalities want to be compensated.

Coleman’s hypothesis that explains the demand for a norm reads: “... interests in a norm arise when an action has similar externalities for a set of others, when markets in rights of control of the action cannot easily be established, and when no single actor can profitably engage in an exchange to gain such rights” (1990b, 42). ‘Similar’ externalities may increase the interest in the emergence of the norm if one is aware that a goal is shared. The second condition can be illustrated with the smoking example. A nonsmoker who suffers from being exposed to a smoker may offer something to the smoker to eliminate the externality. If the smoker agrees a ‘right of control’ has been purchased. If the smoker after the contract still smokes sanctioning is legitimate. If such ‘markets in rights cannot easily be established’ then people demand a norm to regulate the respective behavior. In regard to smoking, imagine you—the reader—offers 3 Euros to a smoker to stop smoking. If the smoker agrees the problem is solved. But such an action is not feasible (i.e. socially accepted): a smoker would probably recommend the nonsmoker who makes such an offer to visit a psychiatrist.⁵

Coleman does not distinguish between different intensities of externalities. It is plausible that many externalities of everyday life are tolerated. For example, there is no norm against crying children in trains or buses or against noise of an occasional party of a neighbor. Only if externalities exceed some threshold the demand for

⁵ If such a market would exist one could easily earn much money: one tries to sit next to a nonsmoker and sells the right to smoke!

a norm emerges. Only then one will utter *commands*, i.e. normative statements such as: one *should* stop smoking or one *must not* violate laws. Expressing such commands is costly especially due to expected counter reactions and complying with commands on the side of the targets of the commands is costly. An individual who expresses such normative demands will only do so if the benefits are relatively high. Thus, only relatively intense externalities will lead to a demand for a norm.

Coleman's second hypothesis that explains the realization of the norm applies the theory of collective action (Olson 1965). This theory addresses the question under what conditions a group of individuals act to achieve common goals. For example, when will the residents of a community act to achieve their common goal of reducing pollution? The common goal is a preference for a public good. This is defined as any good that, if it is provided, can be consumed by every member of a group, regardless of his or her contribution. There is thus no possibility of exclusion or, as Olson puts it, exclusion is not 'feasible' (i.e. is very costly).

This theory can be applied to explain norms because an existing norm is a public good (if it holds it holds for all members of a group). Sanctioning contributes to the provision of a public good: if a non-contributor is sanctioned this increases the likelihood of his or her contribution and, thus the provision of the public good. Externalities are also public bads (pollution) or public goods (the house owner cleans the sidewalk).

The *primary interest* of actors is directed at influencing the externalities. This is thus the *first-order public good*. Norms, backed up by sanctions, are instrumental for the provision of the first-order public good. Norms (or institutions) are therefore called *second-order public goods*.

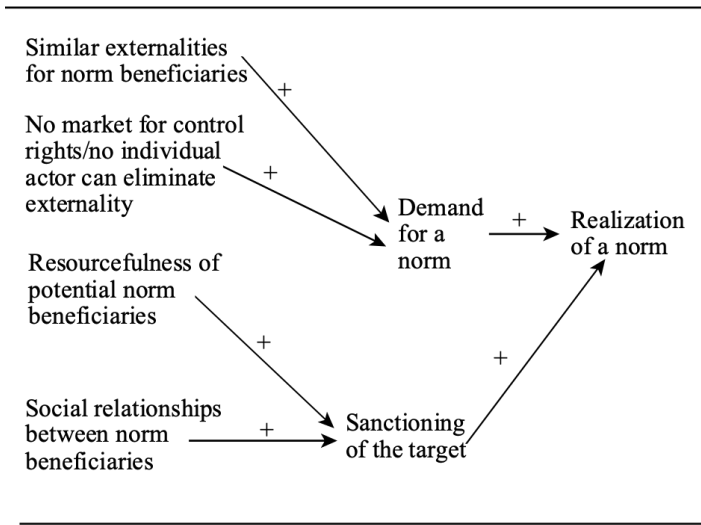
The *contributions* to the provision of the *first-order public goods* vary to a great extent. They may be participation in protests to bring about a law against climate change. Contribution may also consist of working in a citizen initiative to show members of the group how water can be saved in times of big heat. A contribution to the *second-order public good* may be to sanction those who do not contribute to the first order public good, and to sanction those who do not sanction the non-contributors.

The problem of contributing to the provision of public goods in *general* is the free-rider problem: if the good is provided nobody can be excluded. There is thus an incentive to let others contribute until the public good is provided, i.e. to be a free rider who does not contribute.

Coleman's idea is that the provision of the *first-order public good* becomes more likely if people contribute to the provision of the *second-order public good*. Coleman's hypothesis reads: "If there is a social relationship between actors, ... then this overcomes the second-order free rider problem" (1990b, 53). Connectedness with others provides more opportunities to sanction jointly, i.e. to encourage

mutual contributions, and, thus, makes the provision of the first order public good more likely (see also Ellickson 1991, 167, who speaks of the importance of close-knit groups).

Coleman's theory can be reconstructed as a causal model (Figure 2). We added the variable 'resourcefulness of potential norm beneficiaries' that is mentioned in passing. Note that Coleman's writing does not include this model, it is a plausible reconstruction.



Taken from Opp 2018: 176.

Fig. 2: A Reconstruction of James S. Coleman's Theory of Norm Emergence

The theory has some weaknesses (Opp 2018a). One is that only sanctioning is explained. It seems plausible that this brings about the 'realization' (i.e. emergence) of a general norm (see Figure 2). The theory explains thus the acceptance of a norm. But it does not explain internalization in the sense that norm violation elicits a bad conscience, guilt, or shame. People who believe one should follow a speed limit will usually not have a bad conscience if they exceed the speed limit, but

someone who causes an accident that kills a child will feel guilt and shame to a high extent. When such feelings are generated is not explained.⁶

It further seems that under certain conditions close social networks are not sufficient to bring about contributions. For example, if the costs of contributions to the first-order public good are very high (such as being member of an underground organizations that wants to overthrow an authoritarian regime) close social networks will probably not be effective. In general, however, it seems that Coleman's theory is a good explanation of norm emergence, compared to other explanations.

Coleman does not specify in detail the *mechanisms* that lead from the independent variables (see Figure 2) to the 'realization' of the norm. One possibility is that a norm is set by a group. A government may seize the opportunity to gain support by issuing a law that reduces or eliminates the negative externalities. There may be further an agreement (i.e. a contract) about a rule after negotiation.

Another possibility is that the norms originate spontaneously (e.g. Opp 2002, see in general 2018b). Assume that there is no law that prohibits smoking. For many individuals it is a negative externality to be exposed to smokers. The nonsmokers have an incentive to avoid this and very often express their discontent toward smokers and demand to stop smoking. For the smokers smoking becomes more and more costly, and they are increasingly faced with negative sanctioning and normative demands that smoking is socially inappropriate. If, in addition, there is a general norm that hurting others is morally unacceptable, and if this norm can justify not imposing smoke on others, an increasing number of smokers will accept a nonsmoking norm. This is a situation in which no one wants to create a social norm; everybody only wants not to be exposed to an externality. But the actions of the single members of a group bring about an unintended effect: the origin of a *general* norm.

There are numerous mechanisms that lead to the emergence of norms. An early example is the explanation of the spontaneous emergence of property rights of the Indians of the Labrador Peninsula by Demsetz (1967, for a discussion see Eggertsson 1990, 250-280). Ellickson's (1991) explanation of norms to deal with straying cattle is another example. In these cases there is a step by step application of the economic model. The spread of a nonsmoking norm mentioned before is just one example for spontaneous norm emergence. Ellickson's research on how a community solves the problem of straying cattle by informal norms is another example of spontaneous norm emergence. Based on Axelrod's simulation study (1984) repeated interactions is one condition for overall cooperation—which is a norm in

⁶ This issue that cannot be addressed here is a psychological topic. See work on learning theory (e.g. Bandura 1977; Mazur 2017). For moral learning see, e.g., Blair 2017.

the sense of regular behavior. Kliemt (2020, 62) suggests a similar mechanism: “a social custom has emerged exclusively on the basis of predictive expectations and opportunity-taking responses to these expectations.” This ‘custom’ refers to regular behavior. The question remains under what conditions behavioral regularities bring about oughtness. If the majority of residents in a neighborhood take the subway every morning to go to work, this is a ‘custom’ or a regular behavior, but no oughtness is involved. When will it become appropriate behavior to take the subway? If this happens then an is-becomes-ought mechanism is given (i.e. a factual behavior becomes an appropriate behavior – see Opp 1982).

A scope condition for Coleman’s theory is that there exist externalities. However, norms emerge also if there are no externalities. For example, it seems that for certain groups there is a ‘must’ to have tattoos or piercings. This is certainly not due to some previous externalities. Another example is imitation of parents by their children, backed by parents’ normative demands to do certain things. There need not be any externalities. We will return to *imitation* as a mechanism of norm emergence when Vostroknutov’s theory of norm emergence is discussed.

8.2 Vostroknutov’s Theory of Social Norms

It is assumed that a player feels more or less *dissatisfaction* with a payoff allocation in a game. Take the prisoner’s dilemma (see Table 1). The options of actors A and B are to cooperate or defect. If both defect, each gets a payoff of 2; if both cooperate the payoff for each is 3. If one cooperates and the other defects the payoffs are 1 and 4. The player who gets 1 is most dissatisfied because he or she could get 4. If both cooperate the payoff is 3 for each. Since each could get 4 there is only little dissatisfaction. Dissatisfactions can be summarized across actors so that an ‘aggregate dissatisfaction’ for an allocation can be computed (for details see Vostroknutov 2020, 25).

Behavioral alternatives of actor A	Behavioral alternatives of actor B	
	Cooperate	Defect
Cooperate	3, 3	1, 4
Defect	4, 1	2, 2

Tab. 1: Example of a Prisoner’s Dilemma Game

The dissatisfaction of a player causes a *normative valence* of a payoff allocation. This is the extent to which a player considers an allocation 'socially appropriate'. Thus, "the most socially appropriate allocation (the norm) is the allocation with the smallest aggregate dissatisfaction" (26). This is cooperation in the prisoner's dilemma: each player could get payoff 3, so that 3, 3 (both cooperate) is just one unit below 4.

It is further assumed that "a game form as an input ... produces a collection of norm-dependent utilities (one for each player) defining how they will behave" (26). This could mean that acting according to the norm provides utilities. The assumption is further that the utilities generate behavior. The author notes (see 26, footnote 22) that the theory applies only to "'small' strategic interactions of several players." Other conditions for action are mentioned and it is argued that various "phenomena can be easily incorporated in our model" (27). But there are no detailed hypotheses specifying which actions are performed under which conditions.

Finally, a "model of punishment" (28) is added because this "seems necessary for maintaining norm compliance" (28, see also Kimbrough/Vostroknutov 2020c). Our previous discussion indicates that this hypothesis is incorrect. For example, norms may be internalized to such an extent that compliance is automatic, and no sanctioning is needed. But assume that norms are not so strong that compliance has the highest utility and all other incentives for norm following are weak. So far Vostroknutov's model consists of the demand for norms, in terms of Coleman's theory. This is, as this theory shows, not sufficient for the 'realization' of norms. The extensive discussion about punishment and deterrence in criminology (e.g. Apel/Nagin 2017) further indicates that a more detailed model of the causes and effects of punishment is needed.

The theory raises several questions. (1) The theory refers to games, but sometimes there are insinuations that it can be applied to situations that are not yet formalized as games or cannot be formalized games. It seems theoretically useful to refer to any social situation in which actors imagine a better situation and are thus dissatisfied with their present situation. (In a personal communication Vostroknutov argues that his theory can be applied if games are not specified.) (2) The theory explicitly refers to material payoff allocations (27). It seems theoretically useful to include nonmaterial payoffs as well. This is conceded when the author mentions dissatisfactions with a low social status (27). These two points refer to an extension of the scope conditions (i.e. the situations in which the theory applies): it is plausible to apply the theory to any kind of situation in which there is some dissatisfaction—be it material or nonmaterial.

Based on the Coleman theory exposed before, the *dissatisfaction-norms hypothesis* is highly problematic for the following reasons.

- (1) Only dissatisfactions are neither a sufficient nor necessary condition for ‘normative valence’. In terms of Coleman’s theory, ‘normative valence’ seems to mean the demand for a norm (i.e. an interest that a general norm exists). It is hardly plausible that we develop for each of our numerous dissatisfactions a normative valence.
- (2) In Coleman’s theory the central variable is *externalities*. This refers to a certain kind of dissatisfaction, namely with the behavior of others (remember: an example for negative externalities was pollution, for positive externalities the house owner who causes utilities for others but wishes to be compensated). Perhaps externalities in interaction situations are the major kind of dissatisfaction that are one possible condition for norms? Other dissatisfactions are not always important. Think of bad weather (this could be formalized as a game against nature, so that individuals are dissatisfied with certain payoff allocations). It must be left as a question for further research whether certain kinds of dissatisfaction are more likely conducive to a demand for norms.
- (3) It was argued before that externalities and, thus, dissatisfactions are only relevant if they exceed a certain threshold and if there are no norms for tolerating them. Coleman argues that a condition for a demand for a norm is that the dissatisfactions are shared. All these conditions need to be added to Vostroknutov’s theory.
- (4) If ‘normative valence’ means a demand for a norm, it is certainly relevant whether there is a market of control, as Coleman argues. This factor is missing in Vostroknutov’s theory. To put it differently, the interest in a norm is only a *mental construct*: people think that something ought to be the case.
- (5) The second step of norm emergence—the ‘realization’ of a norm—is missing in Vostroknutov’s theory. In Coleman’s theory social networks as possible causes of sanctioning bring about the norm. All this should be added to Vostroknutov’s theory.
- (6) This theory shares the following problems of Coleman’s theory: it is not discussed what the *mechanisms* of norm emergence are (do norms originate spontaneously—and, if so, in what way?—or by design?). Norms also come about without dissatisfaction. Such processes are not discussed in the dissatisfactions-norms theory.

8.3 Extending the Dissatisfaction-Norms Proposition: Norm Emergence through Following Descriptive and Injunctive Norms

One of the mechanisms that bring about norms has not been mentioned so far. It is not due to dissatisfaction. The idea is this: if many other people do something or did something in the past, individuals think this is right. In a similar vein, if many others think something is right, this is accepted. What many others do is, by definition, a *descriptive norm*, whereas what most others think one ought to do is, by definition, an *injunctive norm* (for details see Vostroknutov 2020, 30-32).⁷

Let us illustrate the hypothesis that descriptive norms are imitated with three examples

- (1) A tourist in a foreign country does not know what the tipping custom is. The tourist has several options. One is simply to give the tip that he or she usually gives in his or her own country. The tourist might further ask others in the country (for example the employees in the hotel lobby), or, third, the tourist may observe when he or she first visits a restaurant what the normal tips are that other guests give (or he or she may ask other guests). The tourist then adopts the tipping norm. But this norm is conditional. In his or her own country the tourist gives the tip that is common there. In this example the descriptive norm clearly matters: the tourist does what most others do.
- (2) Assume a taxpayer does not know how many others evade taxes to what extent. He or she calls his tax administration who should know and asks what the rate of tax evasion is. Assume he receives an answer. The taxpayer then thinks this rate of tax fraud should be adopted. Will this always be the effect of the descriptive norm? There will certainly be many 'honest' taxpayers who will not follow the descriptive norm. The internalized norm that forbids tax evasion and contradicts the descriptive norm matters. Furthermore, this might be reinforced by the fear of sanctions.
- (3) A catholic priest has read in the media that numerous colleagues have abused children. The priest computes the number of abused children per priest and thinks this is the morally acceptable number of children that a catholic priest should abuse. In general, however, there will be no imitation. Media reports about the suffering of the victims will rather deter from abuse. Strong internal-

⁷ There are two unpublished papers by Kimbrough/Vostroknutov about descriptive and injunctive norms (2020b and 2020c). For limitations of space it is not possible to discuss these papers in detail. We focus on the article by Vostroknutov 2020.

ized norms against child abuse and expected punishment will further deter from abusing children.

These examples show that a general proposition claiming that a descriptive norm (i.e. the perceived behavior of most others) leads to the formation of new norm is simply wrong. The conditions that bring about such norms are not specified in the work of Kimbrough/Vostroknutov (2020, 30-32).

To extend their work a first suggestion is to explore the widely discussed social psychological research by R. Cialdini and collaborators (see, e.g., Cialdini et al. 1990; Cialdini 2012). The authors conduct experiments that manipulate conditions for matching the behavior of others. One special feature of their work is the idea which is later elaborated in dual-process theories that an important condition for the effects of the behavior of others is the extent to which norms are *activated* (i.e. focused upon).

Kimbrough and Vostroknutov should further consider integrating work about *imitation by rational choice theorists* (e.g., Banerjee 1992; Hedström 1998; Keuschnigg 2015a; 2015b; Király/Oláh 2018, in general see Fryling et al. 2011). Social learning theory (e.g. Bandura 1977) is pertinent as well. The basic idea of this work is that others' *behavior is imitated if individuals believe that imitation yields a relatively high SEU (in terms of VET)*. What these incentives are needs to be determined empirically. To illustrate with the previous examples, our tourists have a high SEU of imitation. Their major incentives could be to satisfy waiters, to avoid negative sanctions of waiters, and to follow an internalized norm to conform to the rules of a foreign country that one visits. The latter is a preexisting incentive. In the tax evasion example, those with a strong compliance norm or with a strong expectation of formal and informal punishment (that are valued very negatively) will not cheat more frequently. Reports on child abuse that describe the long-term suffering of the victims might deter even those who consider this crime. They might further expect stronger formal sanctions due to the effects of those reports on parents and educators: they will pay more attention to the behavior of priests (and other educators) which increases the expected negative sanctions to potential child abusers.

The work of Kimbrough and Vostroknutov excels in formal modeling that is rarely found in sociological work. The models should, nonetheless, be extended by including ideas from sociology and other social sciences some of which have been sketched before.

9 Conclusion and Discussion

This article focused on the contribution of Vostroknutov (2020), that is based on joint work with Erik O. Kimbrough, to the special issue of A&K about norms. In general we share the theoretical approach of the authors to explain normative phenomena: they apply a ‘wide’ version of rational choice theory. The issues we discussed refer to the application of the model. What is largely missing is the consideration of the relevant noneconomic, in particular the sociological rational choice literature. Vostroknutov and Kimbrough’s work is certainly *not* “a ... synthesis of economics approach with psychology, sociology” (2020, 3). Our discussion indicates in detail how their work could and should be extended.

The authors might perhaps also be stimulated by experiments about norms in sociology (mostly by rational choice sociologists). Examples are Berger/Hevenstone (2016); Diekmann/Przepiorka (2015); Diekmann et al. (2015); Keuschnigg/Wolbring (2015); Kroher/Wolbring (2015). There is also an experiment by the economists Balafoutas/Nikiforakis (2012).

For the further development of theory and research about social norms a *cooperation* (not a division of labor) of economists and sociologists (especially rational choice sociologists) seems fruitful. Sociology could contribute substantive propositions to many concrete issues. Furthermore, the extensive literature on sociological methods of empirical research should be taken account of to a greater extent by economists. Economists could bring in their capabilities of formalization.

Especially the contribution of Görges/Nosenzo (2020) shows the fruitfulness *and necessity* of cooperation between economists and empirically oriented sociologists. The authors write about ‘empirical tools to measure social norms’ that are developed by economists in recent years. The authors completely ignore the extensive sociological literature in which norms are measured (for some references see Opp 2015). One would have expected that it is shown what is new in the measurements developed *by economists*, and whether these measurements are superior to the measurements in sociology. The factorial survey seems to be unknown. The authors focus on ‘injunctive norms’ which refer to what most others think one should do. We have discussed this variable in the previous section. The authors ignore the entire non-economic literature about the effects of the social context and the conditions for following others’ behavior or norms, i.e. the causes of imitation (see the references mentioned before).

Acknowledgment: I am grateful to Alexander Vostroknutov for valuable comments on a former version of this paper.

References

- Acemoglu, D./J. Robinson (2012), *Why Nations Fail. The Origins of Power, Prosperity, and Poverty*, New York
- Apel, R./D.S. Nagin (2017), Perceptual Deterrence, in: W. Bernasco/H. Elffers/J.-L. Van Gelder (eds.), *The Oxford Handbook of Offender Decision Making*, Oxford, 121-140
- Auspurg, K./T. Hinz. (2015), *Factorial Survey Experiments*, Thousand Oaks
- Axelrod, R. (1984), *The Evolution of Cooperation*, New York
- Azar, O. H. (2020), The Economics of Tipping, in: *Journal of Economic Perspectives* 34, 215-236
- Balafoutas, L./N. Nikiforakis (2012), Norm Enforcement in the City: A Natural Field Experiment, in: *European Economic Review* 56, 1773-1785
- Bandura, A. (1977), *Social Learning Theory*, Englewood Cliffs
- Banerjee, A. V. (1992), A Simple Model of Herd Behavior, in: *The Quarterly Journal of Economics* 107, 797-817
- Berger, J./D. Hevenstone (2016), Norm Enforcement in the City Revisited: An International Field Experiment of Altruistic Punishment, Norm Maintenance, and Broken Windows, in: *Rationality & Society* 28, 299-321
- Blair, R. J. R. (2017), Emotion-Based Learning Systems and the Development of Morality, in: *Cognition* 167, 38-45
- Boyd, R./ P.J. Richerson (2005), *The Origin and Evolution of Culture*, Oxford
- Cialdini, R. B. (2012), The Focus Theory of Normative Conduct, in: P. A. M. Van Lange/A. W. Kruglanski/E. T. Higgins (eds.), *Handbook of Theories of Social Psychology*, Volume 2, London, 295-312
- /R. R. Reno/C. A. Kallgren (1990), A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places, in: *Journal of Personality and Social Psychology* 58, 1015-1026
- Coleman, J. S. (1990a), *Foundations of Social Theory*, Cambridge/MA-London
- (1990b), The Emergence of Norms, in: M. Hechter/K.-D. Opp/R. Wippler (eds.), *Social Institutions. Their Emergence, Maintenance and Effects*, New York, 35-60
- Demsetz, H. (1967), Toward a Theory of Property Rights, in: *American Economic Review* 57, 347-359
- Diekmann, A./W. Przepiorka (2015), 'Take One for the Team!': Individual Heterogeneity and the Emergence of Latent Norms in a Volunteer's Dilemma, in: *Social Forces* 94, 1309-1333
- /W. Przepiorka/H. Rauhut (2015), Lifting the Veil of Ignorance: An Experiment on the Contagiousness of Norm Violations, in: *Rationality & Society* 27, 309-333
- Eggertsson, T. (1990), *Economic Behavior and Institutions*, Cambridge
- Ellickson, R. C. (1991), *Order without Law. How Neighbors Settle Disputes*, Cambridge
- Elster, J. (1989), Social Norms and Economic Theory, in: *Journal of Economic Perspectives* 3, 99-117
- (1991), Rationality and Social Norms, in: *Archives Européennes de Sociologie* 32, 109-121
- Eriksson, L. (2019), Rational Reconstructions and the Question of Function, in: *Rationality & Society* 31, 409-431
- Erkut, H. (2020), Incentivized Measurement of Social Norms Using Coordination Games, in: *Analyse & Kritik* 42, 97-106
- Feather, N. T. (ed.) (1982), *Expectations and Actions: Expectancy-Value Models*, in: *Psychology*, Hillsdale/NJ

- (1990). Bridging the Gap between Values and Actions. Recent Applications of the Expectancy-Value Model, in: E. T. Higgins/R. M. Sorrentino (eds.), *Handbook of Motivation and Cognition. Foundations of Social Behavior*, volume 2, New York, 151-192
- Fryling, M. J./ Johnston, C./ Hayes, L. J. (2011). Understanding Observational Learning: An Interbehavioral Approach, in: *The Analysis of Verbal Behavior* 27, 191-203
- Gächter, S./E. Fehr (1997), Social Norms as a Social Exchange, in: *Swiss Journal of Economics and Statistics* 133, 275-292
- Gadenne, V. (2011), Über die Validität von Experimenten in der Ökonomie, in: V. Gadenne/R. Neck (eds.), *Philosophie und Wirtschaftswissenschaft*, Tübingen, 51-64
- Görges, L./D.Nosenzo (2020). Using Experiments to Measure Social Norms: Why It Is Important and How to Do It, in: *Analyse & Kritik* 42, 285–311
- Greif, A./D.D. Laitin (2004), A Theory of Endogenous Institutional Change, in: *American Political Science Review* 98, 633-652
- Hedström, P. (1998). Rational Imitation, in: P. Hedström/R. Swedberg (eds.), *Social Mechanisms. An Analytical Approach to Social Theory*, Cambridge, 306-327
- /R. Swedberg (eds.) (1998), *Social Mechanisms. An Analytical Approach to Social Theory*, Cambridge
- Hempel, C. G. (1959), The Logic of Functional Analysis, in: L. Gross (ed.), *Symposium on Sociological Theory*, Evanston, Ill., 271-307
- Homans, G. C. (1967), *The Nature of Social Science*, New York.
- Jasso, G. (2020), Factorial Survey, in: P. Atkinson/S. Delamont/A. Cernat/J. W. Sakshaug/R. A. Williams (eds.), *Sage Research Methods Foundations*, London (forthcoming)
- /K.-D. Opp (1997), Probing the Character of Norms: A Factorial Survey Analysis of the Norms of Political Action, in: *American Sociological Review* 62, 947-964
- Kerr, N. L. (1995), Norms in Social Dilemmas, in: D. A. Schroeder (ed.), *Social Dilemmas: Perspectives on Individual Groups*, Westport/Conn, 31-47
- Keuschnigg, M. (2015a), Imitation und Konformität, in: N. Braun/N. J. Saam (eds.), *Handbuch Modellbildung und Simulation*, Wiesbaden, 903-934
- (2015b), Product Success in Cultural Markets: The Mediating Role of Familiarity, Peers, and Experts, in: *Poetics* 51, 17-36
- /T. Wolbring (2015), Disorder, Social Capital, and Norm Violation: Three Field Experiments on the Broken Windows Thesis, in: *Rationality & Society* 27, 96-126
- Kimbrough, E. O./A. Vostroknutov (2018), A Portable Method of Eliciting Respect for Social Norms, in: *Economic Letters* 168, 147-150
- /— (2020b), *A Theory of Descriptive Norms*, Draft, Chapman University and Maastricht University
- /— (2020c), *A Theory of Injunctive Norms*, Draft, Chapman University and Maastricht University
- Király, I./K. Oláh (2018), Rational Imitation, in: J. Vonk/T. Shackelford (eds.), *Encyclopedia of Animal Cognition and Behavior*, Cham, 1-8 (online)
- Kliemt, H. (2020), Economic and Sociological Accounts of Social Norms, in: *Analyse & Kritik* 42, 41-96
- Kroher, M./T. Wolbring (2015), Social Control, Social Learning, and Cheating: Evidence from Lab and Online Experiments on Dishonesty, in: *Social Science Research* 53, 311-324
- Lewin, P. (2002), The Current State of the Debate Involving The Economics of QWERTY, in: P. Lewin (ed.), *The Economics of Qwerty. History, Theory and Policy*, Essays by Stan J. Liebowitz and Stephen E. Margolis, Houndmills, 244-259
- Mazur, J. E. (2017) *Learning and Behavior* 8th ed., New York
- Merton, R. K. (1957), *Social Theory and Social Structure* 2nd ed., Glencoe/Ill

- Nagel, E. (1956), Formalization of Functionalism, in: E. Nagel (ed.), *Logic Without Metaphysics*, Glencoe/Ill, 262-282
- North, D. C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge
- /R.P. Thomas (1973), *The Rise of the Western World, A New Economic History*, Cambridge
- Olson, M. (1965), *The Logic of Collective Action*, Cambridge
- Opp, K.-D. (1982), The Evolutionary Emergence of Norms, in: *British Journal of Social Psychology* 21, 139-149
- (1999), Contending Conceptions of the Theory of Rational Action, in: *Journal of Theoretical Politics* 11, 171-202
- (2002), When Do Norms Emerge by Human Design and When by the Unintended Consequences of Human Action? The Example of the No-Smoking Norm, in: *Rationality & Society* 14, 131-158
- (2005), Explanations by Mechanisms in the Social Sciences. Problems, Advantages and Alternatives, in: *Mind & Society* 4, 163-178
- (2013a), Norms and Rationality. Is Moral Behavior a Form of Rational Action?, in: *Theory & Decision* 74, 383-409
- (2013b), What is Analytical Sociology? Strengths and Weaknesses of a New Sociological Research Program, in: *Social Science Information* 52, 329-360
- (2015), Norms, in: J. D. Wright (ed.), *International Encyclopedia of Social and Behavioral Sciences*, volume 17 (2nd edition), Oxford, 5-10
- (2018a), Externalities, Social Networks, and the Emergence of Norms: A Critical Analysis and Extension of James Coleman's Theory, in: *Social Research* 85, 167-196
- (2018b), The Interdependence of Spontaneous Order and Institutional Design, Table Manners, Language, Daylight Saving Time and the Erosion of Institutional Design under Communist Rule, in: A. Mica/ K. M. Wyrzykowska/ R. Wiśniewski/ I. Zielińska (eds.), *Sociology and the Invisible Hand*, Frankfurt, 197-228
- (2020a, forthcoming), Rational Choice Theory and Methodological Individualism, in: P. Kivisto (ed.), *The Cambridge Handbook of Social Theory*, Cambridge
- (2020b, forthcoming), Are Individuals Utility Maximizers? Empirical Evidence and Possible Alternative Decision Algorithms, in: A. Mays/A. Dingelstedt/V. Hambauer/S. Schlosser/F. Berens/J. Leibold (eds.), *Grundlagen-Methoden-Anwendungen in den Sozialwissenschaften*, Festschrift für Steffen M. Kühnel, Wiesbaden, 421-440
- Popper, K. R. (1959), *The Logic of Scientific Discovery*, New York
- Stigler, G. J. (1950a), The Development of Utility Theory I, in: *Journal of Political Economy* 58, 307-327
- (1950b), The Development of Utility Theory II, in: *Journal of Political Economy* 58, 373-396
- Stroebe, W./Gadonne, V./B.A. Nijstad (2018), Do Our Psychological Laws Apply Only to Undergraduate Students? External Validity Revisited, in: *Basic and Applied Social Psychology* 40, 384-395.
- Vostroknutov, A. (2020), Social Norms in Experimental Economics: Towards a Unified Theory of Normative Decision Making, in: *Analyse & Kritik* 42, 3-39
- Wigfield, A./Tonks, S./S.L. Klauda (2016), Expectancy-Value Theory, in: K. R. Wentzel/D. B. Miele (eds.), *Handbook of Motivation at School*, London-New York, 55-75