CHRISTOPH LUMER

The Content of Originally Intrinsic Desires and of Intrinsic Motivation

(1) Originally intrinsic desires are motivational intrinsic desires for those aims for which the subject has no further reason and, since one may have forgotten the reason, never had any reason. (2) The most familiar and important hypothesis about these desires is psychological hedonism. Some examples of the strong explanatory force of psychological hedonism are given. (3) But there are originally intrinsic desires beyond hedonic ones, namely feeling-induced desires, especially emotion-induced desires, e.g. when in a rage the desire to destroy or hurt the “aggressor”. The origins and mechanisms of such desires are expounded and their irreducibility to hedonic desires is shown. (4) A comparison with other theories shows the progress being made by this approach. (5) Being instable over time feeling-induced motives are not apt as a basis for rational utility functions.

1. What Are Originally Intrinsic Desires?

The aim of this paper is to explore the empirical content of originally intrinsic desires: What are the originally final aims people are striving for? But before being able to answer this question some clarification of its meaning is useful.

I presuppose some vague form of desire-belief-theory or expectancy-value-theory of action: If actions are based on deliberation — even on a very reduced form of deliberation — the content of such deliberation (among other things) is this: The value or desirability of consequences of the action or of some set of alternative actions are considered, weighed or measured. Any time when during such deliberation the value or desirability of some consequence or of the action itself is established I call this a “motivational value judgement” and the resulting attitude a “(positive or negative) desire” — knowing that this is technical jargon

Acta Analytica 18, 1997; S. 107-121
and not ordinary language. These desires may be rather peripheral or implicit in the sense that one need not make an explicit or verbal value judgement. E.g. one simply knows that being hurt painfully is bad; one must not say that internally; but if asked one could say that it is bad.

Desires and motivational value judgements may be intrinsic, extrinsic, total or expectancy desires. In intrinsic desires a certain state of affairs is valued in a certain way for its own sake, not for its consequences or other states of affairs implied by it. If subjects are asked why they (intrinsically) value this state of affairs in a certain way they cannot give any further answer: 'There is no further reason or consequence why this is good (or bad); it is simply good, good in itself.' All the other desires are dependent on intrinsic desires. That does not mean that people valuing something only as totally good in that moment must be aware of the intrinsic desire behind that value judgement. But on reflection they may give you their reasons. Probably most aims people are consciously striving for (rushing to work, having breakfast, washing their hands, earning money etc.) are not intrinsically desired. Instead such aims are rather directly accessible by well known actions, and they are held to have a certain positive total desirability because their realisation does ensure some intrinsic ends.

One way of introducing intrinsic desires has been this: People can be asked for their reasons why they desire something; and when they cannot give any further reason this must be an intrinsic desire. But sometimes people forget their further reasons; they forget that originally, when acquiring this desire, they had an intrinsic desire which they thought would be fulfilled by the realisation of that desire. This kind of forgetting or even repressing may happen especially when the original aim was only to prevent some punishment or when there never was a clear reason why this state of affairs should be good but there was only the assertion by some authority that this was good for the subject. This paper is not interested in intrinsic or supposed to be intrinsic desires in general (according to the mechanisms just explained their content may be of a theoretically unmanageable variety) but only in originally intrinsic desires.

2 Psychological Hedonism

Psychological hedonism is the thesis that all sorts of feelings of the agent, namely pleasant or unpleasant bodily feelings, emotions and moods, are the objects of his originally intrinsic desires and that they are valued according to the integral of their intensity. Psychological hedonism is the most important and best known hypothesis about the content of intrinsic desires. Strong psychological hedonism claims that the agent's feelings are exactly the objects of his originally intrinsic desires; weak psychological hedonism holds that beyond the agent's feelings there may be other objects in his originally intrinsic desirability function.

Weak psychological hedonism seems to be a certainty but one may wonder if strong psychological hedonism is true. Psychological hedonism very often has shown its strong explanatory force. And I would like to add two further proofs of this force:

1. Psychologists speak of "intrinsic motivation" in a variety of meanings. But the most common meaning seems to be: autotelic motivation, i.e. that someone is doing something with a motivation not aiming beyond the immediate situation: playing, humming, working with flow. On first impression the definition of this type of "intrinsic motivation" suggests that in these cases the respective activities are judged as intrinsically good. But this is false: In all these cases people immediately get a deep satisfaction, i.e. a positive feeling, out of their activity; and even if people follow other ends in this situation (reaching the goal of their game or solving the problem) and are not or only very little attentive to their feelings it seems to be clear that the whole mechanism of autotelic activity would not work if there was no satisfaction: People give a high desirability to the respective

2 That even this impression is false and that the hedonic desirability function has to be corrected a bit (for handling cases like that of Nozick's experience machine (Nozick 1989, ch. 10)) I have shown in: Lumer 1996.
4 Cf. the contributions of Rheinberg, Schneider and Sokolowski in: Kuhl / Heckhausen 1996. — In Deci's and Ryan's well known theory "intrinsic motivation" seems to mean (the authors obviously have some difficulties to define what they mean) something like: motivation directed at immediately following experiences but not experiences deriving from the satisfaction of bodily needs, or in short: motivation directed at immediately following emotions (cf. Deci / Ryan 1985, 11). So the phenomena Deci and Ryan want to inquire may be the same as those others have in mind when they speak of "autotelic motivation".

1 For a discussion of the different meanings of "desire" see: Marks 1986; Schueler 1995.
activities because among others these activities grant that type of satisfaction. This clearly can be seen during that phase of their activity when people get fascinated by it: In the beginning the scope may be rather arbitrary and subjects may start the respective activity with little enthusiasm; but little by little they detect the challenging features, find their satisfaction in mastering them, and the interest in that activity grows enormously — which means: The subjects now, at least implicitly, have detected the positive hedonic consequences of that activity and consequently change their valuation of it. So even in so called “autotelic behavior” the intrinsically desired aim is a positive feeling, and this case is covered by psychological hedonism.

2. Another example for the explanatory force of psychological hedonism are the implicit motives of McClelland and coworkers (McClelland / Koestner / Weinberger 1989): Implicit motives are attributed to subjects according to the results of a thematic interpretation test: The subjects have to write down or tell their thoughts inspired by picture stories; according to the central concerns in these thoughts motives like affiliation, achievement or power are attributed to the subjects. Motives attributed in this way have a rather high prognostic value with respect to the general type of activity and the life project these people follow. But these motives do not correlate significantly with self-attributed motives; therefore they are termed “implicit motives”. The explanation of this discrepancy is this: People with a certain implicit motive act only as if they had that motive; actually they have no such motive but they get outstanding pleasure from doing things someone does who has such explicit motives. People with a high implicit achievement motive like to do jobs well; people with a high implicit power motive like to command other people and structure social situations etc. The origin of this attitude seems to be that they rather early in their life have acquired firstly the cognitive skill to detect features of a situation which might give satisfaction by power, feelings of self-esteem or of social harmony etc. and secondly a scheme of attention which grants that during performance they register especially such features of the situation the realization of which triggers the emotions of power feelings, feelings of self-esteem, or of social harmony respectively. — So in this rather strange group of motives, too, the actually central and intrinsic desire is a hedonic one: to have certain positive emotions.

Footnote: In the following I depart a bit from the explanation of McClelland et al.

3. Beyond Psychological Hedonism: Feeling-Induced Intrinsic Desires

In spite of the strong explanatory force of psychological hedonism I think there are originally intrinsic desires which are not hedonic; I call them “feeling-induced desires/motives”. In working out the explanatory force of psychological hedonism one risks overstretching the hedonistic explanation scheme always suspecting that the subject might have in mind — at least originally — some hedonic consequences of this type of action. Therefore, for explaining my thesis I have chosen a paradigm case where such hedonistic reduction seems to be unrealistic or far fetched: Think of a small girl of three or four years who has been provoked by her elder brother (of ten years age), and being really furious with him she takes his arm and bites him with all her power in his forearm directly above his wrist-watch. Surely our girl knows what she is doing — biting her brothers forearm — and she will have known in advance that this action will “damage” or “destroy” the aggressor, perhaps even that this will injure him and that he will suffer. And she has chosen well the point where to bite him: in a place where her mouth has enough grip, and not on the wrist-watch which would hurt her. Damaging or destroying the aggressor seems to be the (desired) aim of that action, and there seems to be no other aim behind that. But that the aggressor is damaged or destroyed is no feeling of the agent so that we have a not hedonic originally intrinsic desire. Afterwards the girl will be satisfied in a crude moral way. But it seems to be too far-fetched to suppose that such a little girl already knows about such hedonic consequences of her acting out of rage; at least when she acts the first times out of rage she cannot have the empirical knowledge about these hedonic consequences, but must acquire this knowledge and perhaps some time, some years later may even have the intrinsic hedonic aim of being morally satisfied.

What is the general mechanism behind this type of intrinsic motive? I have called such motives “feeling-induced”; and with “feeling” I mean bodily feelings, emotions or moods. Let us first look for the emotion-induced motives! Emotions themselves typically — but not always — arise out of some affective valuation of a situation or thought: The subject considers some situation or has some thought and classifies this in a specific way, e.g. as proof of his own strength or as an (unjustified) aggression of somebody else against himself (with the other neither being too strong nor too insignificant). This classification implicitly is a valuation, i.e. the classificatory attributes are meant to be positive or negative attributes.
Table 1: Succession of events belonging to emotion-induced desires

The table lists a complete sequence from the first emotion to the accompanying emotion. This sequence may, of course, be interrupted after nearly every step.

I. First emotion A

1. Thoughts or consideration of a situation
2. Affective valuation of that thought or situation (i.e. classification as being of a certain (positive or negative) type G)
3. Emotion A (i.e. emotional feelings)
4. Accompanying bodily phenomena

II. Emotion-induced-desire and action

5. Consideration of some action a and of its consequences
6. Classification of one of the consequences c as F (F being the affective value criterion of the satisfying emotion B belonging to A)
7. Emotion-induced desire, i.e. intrinsic motivational appraisal of the consequence’s c being F (as positive according to the criterion of emotion B, with the absolute value being proportional to the intensity of Emotion A)
8. Total appraisal of action a as being optimum
9. Action a
10. Occuring of consequence c

III. Satisfying Emotion B

11. Perception of c
12. Affective valuation of c as being F (and, therefore, being positive)
13. Emotion B
14. Accompanying bodily phenomena

The next step is that the affective valuation causes the appertaining emotion, e.g. pride or rage; and the content of that valuation is the propositional content of the emotion (cf. table 1, steps 1-4). Every type of emotion has its specific affective valuation like pride and rage (cf. above and table 1, step 2), e.g. fear rests on the classification and valuation that something rather harmful probably will happen (cf. Solomon 1993, 220-310). Such classifications and valuations are not hedonic; but they are only affective in the sense that they cause certain emotions; they are not motivational. So this is not yet the not hedonic originally intrinsic motivational desirability function we are looking for. But every type of emotion is linked with another, satisfying type of emotion in the way that tokens of that emotion seem to aim at tokens of this emotion. Rage has (moral) satisfaction as its satisfying companion, happiness has attachment as its satisfying companion, fear has relief as its satisfying companion etc. This sort of companionship that every emotion is aiming at another emotion is not ordered in a circular way but points to ultimate emotions having themselves as their satisfying companion. Such ultimate emotions are always some form of satisfaction. The nonhedonic motivational and originally intrinsic valuation now is this:

Law of emotion-induced desires:

1. If someone has a certain emotion A and
2. during his emotional arousal classifies some consequence c of a potential action of himself as F and
3. F is the classificatory attribute of the (positive) affective valuations of the satisfying emotion B belonging to A then
4. the state of affairs that c is F is desired (originally) intrinsically proportional to the strength of the emotion A.

Consider our little girl: The girl is furious with her brother (emotion A; cf. table 1, steps 1-4); the satisfying accompanying emotion is (moral) satisfaction (B); and the affective valuation causing this (moral) satisfaction rests on classifying that some villain or aggressor has got his just punishment or, a bit more primitive, is damaged or destroyed (F); now the girl considers the possibility that one of her

---

6 Some further examples are: To satisfaction, envy and curiosity belongs satisfaction; to indignation belongs gratification; to disappointment / frustration and hope belongs joy; to relief belongs a feeling of security; to joy, happiness, negative and positive empathy belongs positive empathy; to negative self-esteem belong — according to the intensity of this emotion — positive self-esteem or gratification (about misfortunes happening to oneself); to shame belongs a feeling of being accepted; to feeling guilty belongs relief; to pride belong positive self-esteem and pride.
actions will (c) damage the aggressor (F; cf. table 1, steps 5-6); this possible consequence then is motivationally (positively) intrinsically desired (cf. table 1, step 7); and the strength of that desire corresponds to the intensity of our girl's rage.

Such originally intrinsic desires are emotionally induced; i.e. being in a certain emotional state is the central and necessary cause of such desires. And if the emotion fades the desire fades too. So the dependency on the inducing emotion is responsible for a characteristic instability of such desires over time. I shall return to this important feature later on.

Emotionally induced desires in a certain sense rest on anticipating affective valuations. The emotion of our girl e.g. rests on a first affective classification and valuation that she has been offended by her brother (cf. table 1, step 2); the content of the emotionally induced desire is instead that it is good that the offending brother will be hurt (cf. table 1, step 7). This content of the emotionally induced desire is the same as that of the (possible) later affective valuation which actually causes satisfaction — apart from a different time index: 'My offending brother has been hurt or damaged.' (cf. table 1, step 12). In this sense the emotionally induced motivational valuation is an anticipating affective valuation with motivational function. One may suspect that this anticipating affective valuation may immediately cause the pertinent emotion, i.e. in our example moral satisfaction, because for arousing emotions mere thoughts of the specific content are sufficient. This actually may happen but only if the subject revels in thoughts of that specific content, e.g. thoughts of revenge. Then the main concern of the subject changes, the first emotion temporarily makes room for the conjoined satisfying emotion, and so the emotionally induced motivation is diminished or even erased. But this is not the case we are considering; in our case there was only one short thought that a certain course of action will damage or hurt the aggressor (table 1, step 6). Under the specific circumstances this is enough for (motivationally) desiring to hurt him (step 7), but it is not enough for causing moral satisfaction.

Until now we have considered only emotionally induced desires. There seems to be a similar phenomenon in bodily feelings which here I only can mention: Having a strong positive bodily feeling at least sometimes people desire the continuation of that feeling more intensely than would be adequate according to the normal anticipatory hedonic desirability function; and having a strong negative bodily feeling they excessively desire its cessation (cf. e.g. Brandt 1979, 39 ff.). So people with high sexual lust might be prepared to pay a much higher price for the continuation of their lust than they would do in advance with a cool head; and people with strong pains analogously might be prepared to pay a much higher price for the cessation of their pains than they would do with a cool head. These overvaluations are feeling-induced too; but they are a bit less interesting than the emotion-induced desires because their content is still hedonistic: prolongation of pleasant and cessation of unpleasant bodily feelings. But they are not normal hedonic motivational desires because they are desires out of proportion. And this e.g. may explain several cases of weakness of will.

Finally, there may even be originally intrinsic desires induced by moods. At least there are some well-known effects which might be interpreted in this way. Firstly, in negative moods people are impatient in the sense that they are less willing to bear any kind of negative feeling: They postpone unpleasant tasks the further the more depressed they are and they prefer smaller but immediate rewards to greater but later rewards. And, secondly, in positive moods people are much more willing to help others (Morris 1989, 100; Dovidio 1984). But one problem with these findings is that it is not always clear that the effect is due to moods and not to emotions. Another problem is that all these findings can be explained in a different way; e.g. the preference for the "smaller" reward may in fact be a preference for a "smaller" reward having now much greater effects in improving the negative mood than the "bigger" reward would have later on. A third problem is that we are always in a certain mood but there cannot be permanent distortion of normal desires. So if the mentioned effects must be explained by mood-induced desires there will be a threshold of mood intensity which must be exceeded for the induction of a desire. — For the moment I have to leave open the question if there are mood-induced originally intrinsic desires.

Feeling-induced desires contribute to decisions like other desires: They imply a certain valuation of certain aspects of actions; and this valuation is one summand for the total valuation of these actions. So feeling-induced desires do not have an automatic priority; they may be outweighed by other desires.

7 Morris (1989, 109 f.) gives a list of confirmations for these effects.
4. What is New about this Theory?

What theoretical progress has been made with the proposed theory? Among living philosophers strong psychological hedonists are a small minority; one of them is Carolyn Morillo (Morillo 1990; 1992). Though her theory makes good distinctions and is one of the psychologically best informed theories one problem with it is that its scope is explicitly delimited to basic motives like hunger (Morillo 1992, 9). But the main problem is that Morillo — as a remnant from behavioristic learning theory — treats intrinsic desires as conceptually equivalent to reward events (cf. e.g. ibid. 17); so intrinsic desires induced by feelings and expressing themselves only in valuations of consequences of actions are already conceptually excluded. (In an expectancy-value-theory there is no need that realized intrinsical values must be experienced as rewards.)

On the other hand there are many philosophers who explicitly or implicitly criticize strong psychological hedonism and advocate that there are further intrinsic desires, e.g.: autonomy (Young 1982), personal expressiveness, i. e. living in accordance with one’s true self (Waterman 1990), virtue and love (Davis 1977), intellectual consistency, psychological maturity and moral rectitude (Nelson/Wilker 1975), perfection of character (Tatarkiewicz 1984, ch. 24), among other things having the right contact to reality, wisdom (Nozick 1989, chs. 10; 11; 23) — to cite only a few of them. I think, except from love (in a special meaning), none of these objects is originally intrinsically desired. These philosophers have not really scrutinized such desires if they are really originally intrinsic desires: Sometimes even the intrinsically good is mixed up with the intrinsically desired, sometimes simple aims are equated with intrinsic desired objects, often a careful examination if there are other (esp. hedonic) desires behind such desires is missing, nobody distinguishes between originally and not originally intrinsic desires etc. But the main problem with these suggestions is that they come from philosophers primarily interested in normative (moral or prudential) questions; therefore, these suggestions remain isolated and are not integrated in a complete theory of intrinsic desires or motives and of desires in general. So in spite of the mentioned publications in contemporary philosophy today there still does not exist any nonhedonistic theory of originally intrinsic desires — in a somewhat more demanding sense of “theory”.

In emotional psychology the idea that emotions cause motivational tendencies is rather common (cf. Izard 1993, 72; Lazarus 1974, 21 f.; Plutchik 1980, 155; Scherer 1990, 5; Tomkins 1970). (Some emotional psychologists even reduce emotions to a changing or induction of motivational tendencies (Frijda 1986, 460; 466; 469; 479; Izard 1981, 63 etc.; Lang 1988, 186). But this reduction eliminates the too familiar phenomena.) But, firstly, the alleged content of the induced motives is rather crude: Often the examples are taken from animal psychology; and the most popular example is that fear induces tendencies to run away. Surely, frightened people often run away because they think they can thus escape danger; but there are dangers where running away does not help; e.g. if I fear having a heart attack I will call for help and then probably sit down. So these theories do not clarify the exact content of our desires stemming from emotions, not even the details of a flight. Secondly, the emotional theories of motivation are not integrated in general theories of motivation. In particular there is no explanation how emotionally induced motives may be combined with other motives. E.g. somebody might be frightened by the prospect of dental treatment, but nonetheless he goes to the dentist because he thinks that this is better for his teeth and for avoiding future pain. In such a case the emotionally induced desire is outweighed by other desires. According to the motivational models of emotional psychology such things are not possible.

---

8 Richard Brandt is another strong psychological hedonist (Brandt 1979, 132; 134), likewise psychologically well informed. With respect to the expansion of the range of intrinsic desires beyond hedonism the main problem in Brandt’s theory is that he too does not systematically investigate the field of possible candidates. — In addition to this his theory is — being a bit older — still rather behavioristic so that even in the positive, hedonistic part of his psychological theory there are big problems: He e.g. defines the ‘pleasing’ motivationally as the tendency to produce certain actions (Brandt 1979, 40) — though feelings and motivation are conceptually separated: There may be people having pleasant feelings and unfortunately being motivated to immediately interrupt such feelings or being unable to act at all. (The last case seems to be really existent in anencephalic babies who seem to have pleasant feelings from tasting sugar but have no “hardware” for being able to act at all (Steiner 1973, esp. 263; cf. Morillo 1992, 20).) A related problem is that Brandt does not distinguish between feelings of e.g. hunger changing rapidly over time and the stability of the valution of such feelings; so he thinks falsely — as many psychologists do — that the valence of eating changes over time (Brandt 1979, 57 f.). But the truth is that eating at different times has different values but these values ideally — remain constant over time. Therefore having just eaten I now can disvalue the fact of being hungry tomorrow morning and can make arrangements now to prevent this unpleasant state. According to Brandt’s theory this would be impossible.
In motivational psychology a host of lists of motives exists, on the one hand alleged to be complete taxonomies (like e.g. that of Maslow 1954) and on the other hand chapters in textbooks (e.g. Heckhausen 1989; McClelland 1985) devoted to the main motives like achievement, help, aggression, affiliation, power etc.; the status of the latter lists is not clear. Such taxonomies and lists may have their special merits but, surely, they are not developed for answering the question of this paper. Lists of motives like achievement, help etc. seem to be more classifications of frequently occurring forms of apparent behavior and traits of character than of fundamental motives; all these motives can be reduced to hedonic motives (even if they are only implicit (cf. above, section 2). The question (what is the content of originally intrinsic desires and motives?) itself for psychologists seems to be rather far away from their own interests and difficult to understand, which may partly be due to the fact that "intrinsic motive / motivation" in psychology has a variety of meanings, the most important of them having little to do with the philosophical meaning explained in section 1.

But there are also motivational psychologists who take up the motivational hypotheses of emotional psychology, Heckhausen perhaps in the most polished form: Besides the usual motives leading to decisions in the expectancy-value-mode he claims that emotions are a rudimentary motivational system with emotionally induced motives leading directly to decisions and actions without any expectancy-value-elaboration (Heckhausen 1989, 71-76, esp. 74). But in this form his theory has nearly all the problems which I have attributed to the motivational hypothesis of emotional psychology: The contents of the emotionally induced motives remain as crude and rough as they are in those theories; and the two types of motives cannot be combined and contribute to one and the same decision.

Compared with all these theories the advantages of my proposal are these: It preserves all the advantages of psychological hedonism but admits further originally intrinsic desires and thus may give satisfaction to those people deeply unsatisfied with strong psychological hedonism. The addition to psychological hedonism is not ad hoc but the core of a real theory. Part of my theory is a couple of hypotheses about the exact content of emotionally induced desires. And the theory integrates the emotionally induced desires in the normal framework of expectancy-value- or desire-belief-theory of decision.

5. Some Theoretical Consequences of the Hypothesis of Feeling-Induced Desires

Feeling-induced desires seem to be an evolutionary older motivational system than the normal hedonic motivational system. And it is a much more primitive motivational system than this one: Firstly, emotionally induced intrinsic aims are rather immediate changes of the exterior situation. This change often will be beneficial for the subject; otherwise this motivational system would not have survived evolution. But there is no room for changing these aims if they are not beneficial for the subject or if there are possibilities of advancing the (longterm) well-being of the subject much more. Secondly, feeling-induced desires are bound to current feelings so that they change rapidly over time. This implies that they are not apt as a basis for long-term planning and long-term decisions: At the moment of a possible long-term decision, e.g. to hurt one month later a probable aggressor who will have revealed himself as such only that month later, the intrinsic desire simply does not yet exist; and if in rage one plans to hurt the aggressor only one week later when there will be a better opportunity for doing so then that week later, when the moment of action has come, often the emotion will be lacking and with it the emotionally induced desire — with the consequence that the agent will decide otherwise. In this respect feeling-induced desires differ sharply from normal hedonic desires: If I know that I might suffer from hunger or anxiety one year later I intrinsically disapprove of such feelings now, one year later or whenever. And the stability of such valuation is the basis for the fact that I now can plan to avoid such feelings which otherwise will arise only very much later.

Emotion-induced desires are philosophically or theoretically interesting because they successfully challenge strong psychological hedonism so that only weak psychological hedonism may be true. A theory of originally intrinsic desires taking feeling-induced desires into account might be acceptable for those with reservations with respect to strong psychological hedonism. Perhaps there are still other not hedonic originally intrinsic desires; but for the moment I do not see any. On the other hand the instability over time of feeling-induced originally intrinsic desires makes them unsuitable as a basis for rational desirability functions: One aim of rational desirability functions is to make possible long-term planning for taking advantage of good opportunities and for cheaper satisfaction of desires in the long run; this is not possible if desirability changes over time (Lumer 1998).
From this point of view quickly changing desirabilities seem to be quite irrational. And for this reason, too, we often regard acts committed in the heat of passion as irrational: Directly after having reached his aim the agent may already regret his deed; this does not look like rational behavior.

References


