OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

WW FACULTY OF ECONOMICS AND MANAGEMENT

Medical Decision Making: Perspectives of Patients and Physicians

Inauguraldissertation

Zur Erlangung des akademischen Grades

Doctor rerum politicarum

vorgelegt und angenommen

an der Fakultät für Wirtschaftswissenschaft

der Otto-von-Guericke-Universität Magdeburg

Verfasser:	Judith Natalie Trarbach
Geburtsdatum und -ort:	03.12.1983 in Dachau

Arbeit eingereicht am:	28.08.2013
Gutachter der Dissertation:	Prof. Dr. Dr. Bodo Vogt
	Prof. Dr. Bernt-Peter Robra
Datum der Disputation:	21.November 2013

Danksagung

Mein Dank gilt der Konrad Adenauer Stiftung, mit deren Unterstützung ich dieses Dissertationsprojekt angegangen bin und die mich bis zu dessen Ende gefördert hat. Natürlich bestand ein wesentlicher Teil dieser Förderung aus finanzieller Unterstützung, jedoch war diese Zeit für mich darüber hinaus sehr wertvoll und geprägt von spannenden Themen, die mich während meiner Dissertation nach links und rechts schauen ließen sowie auch von neuen Bekanntschaften zu interessanten Menschen. Ich bin dankbar, dass ich die Chance bekommen habe, Teil des lebendigen Netzwerks der Stiftung zu werden.

In Prof. Vogt habe ich einen Betreuer gefunden, der von Anfang an beständig an mich und meine Stärken geglaubt hat und sich nicht scheute, eine Psychologin in Gesundheitsökonomie promovieren zu lassen. Diese Chance habe ich sehr gern ergriffen und ich freue mich über den enormen Wissenszuwachs zu dem er mich gebracht hat. Seine Unterstützung sowie die Freiheit, meine Projekte eigenständig durchzuführen, haben meine Promotion zu einer sehr spannenden Zeit für mich gemacht.

Danken möchte ich auch Prof. Dr. Robra, der immer Zeit für mich hatte, wenn ich ihn um einen Austausch gebeten habe, der mich so offen integriert hat und dem ich so gern zugehört habe, wenn er sich die Zeit nahm, mich an seinen Themen und Gedanken teilhaben zu lassen.

Darüber hinaus bin ich froh über die Zusammenarbeit mit den Kollegen an meinem Lehrstuhl sowie aus dem Institut für Mikro- und Sensorsysteme. Besonders Prof. Dr. Lucklum und Stephan Adler haben weder Zeit noch Mühen gescheut, mich in der Durchführung meiner Experimente durch die Nutzung ihrer Räume und Technik zu unterstützen.

In allem begleitet hat mich mein Mann Thomas, der zu jeder Zeit meine Arbeit mit Zutrauen, Geduld und Verständnis unterstützt hat.

Table of Contents

List of Figures

List of Tables

Introduction

During the last 30 years, the amount of research studies conducted in the context of health care has augmented massively. In fact, various disciplines are engaged in these investigations. Among these are e.g. economics, medicine, psychology, sociology, politics or philosophy. The core question for all of these disciplines asks how we can ensure the best health care service for everybody given limited resources. This question is the basis for a wide variety of subordinated problems, be it in terms of the core bundle of essential health care services, in fair allocation of resources, in the assignment of responsibilities in the health system or in ethical discussions on the aim of health care.

This dissertation sets the focus on the two central actors: care giver and care receiver – physician and patient. It combines methods and theories from different disciplines i. e. behavioral economics, psychology and health economics. Within three connected studies, decision making behavior of physicians and preferences of patients are investigated.

1. The first study focuses on resource allocation decisions a general practitioner faces every day. Often, financial or temporal resources are insufficient for all individuals in need. Accordingly, one has to decide who is treated how and when. Possible decisive attributes are investigated which the physician uses to come to her allocation decision.

2. The second study focuses on specific situations of patients excluding the care giver. It demonstrates the relevance of treatment attributes which cause individuals to prefer one treatment over the other. This study simulates the situation of patients being involved in the decision to choose one out of two possible therapies for their disease.

3. The third investigation deals with the common approach to measure changes in quality of life, the QALY-concept. This concept is necessary when it comes to evaluations of treatments or decisions on appropriate resource allocation. We challenge a theoretical assumption on risk attitude which is included in the concept in terms of risk preferences of the affected patients.

Within these three studies, a special focus lies on the methodology. All investigations are realized with an experiment. The aim is to observe real decision behavior either on the side of patients or physicians. Experimenters distinguish between experiments with hypothetical and real incentives. In experiments with hypothetical incentives, subjects receive a flat fee for participation. This is different for experiments with real monetary incentives where decisions influence the individual pay-off (Hertwig & Ortmann, 2001). Repeatedly, differences in individual decision making have been observed. Their outcome depends on the consequences of decisions as either real or hypothetical (Blumenschein et al., 2001; Pesheva et al., 2011). Although rationality violations often do not disappear when real incentives are introduced, real incentives reduce presentation effects compared to hypothetical incentives (Camerer & Hogarth, 1999). Such presentation effects might result in one of the following three alternatives. Risk aversion increases, if lottery choices are real and not hypothetical (Holt & Laury, 2002; Harrison & Rutström, 2008) (1). Generosity decreases, if decisions result in monetary instead of hypothetical incentives (2). However, this last effect might be attributed to the height of stakes. In hypothetical experiments, stakes are typically high. Experiments with stakes at the same height typically show no difference (Kühberger et al., 2002). Finally, humans tend to overestimate the theoretical value of goods compared to the value they are willing to pay (Murphy et al., 2005) (3). Because of these behavioral observations, the studies presented in this work all avoid hypothetical questioning. Instead they include two types of real consequences described in more detail subsequently.

Consequences for the individual well-being

Decision making in the health care context always refers to changes of health, either an improvement or a deterioration of the quality of life for the affected patient. Within an artificial context such as an experimental setting the simulation of such a condition is difficult. The inclusion of real patients who suffer from a bad health state awaiting treatment which finally might not be delivered as the consequence from the experiment is unethical and can never be an option. As a solution in the experimental context, healthy students become patients during the time in the laboratory. Therefore, their level of well-being is reduced artificially by means of pain induction. The experiments base on the assumption that almost every reduced level of well-being includes pain sensation. Accordingly, pain induction is a quite general aspect of bad health and representative for many conditions. In this work, a standard procedure is implemented which is used in medical research for about 80 years in different contexts: the cold pressor test (CPT) (Hines & Brown, 1936). In this test a subject has to immerse a hand into cold water without moving it for a given amount of time. The test endures several seconds or minutes. In general, a pain sensation occurs within water colder than 18° Celsius. Consequently, in most studies water of about 4° Celsius is used. The pain is deep and constant for several minutes and its intensity relates to the selected temperature (Chéry-Croze, 1983). Moreover, it is described in the literature to be comparable with chronic health states (Mitchell et al., 2004). This pain induction procedure is not affecting general health at all and is even recommended for usage with children (Birnie et al., 2011). After arriving in the laboratory, every participating subject was informed explicitly about the test and signed a consent form to accept the pain induction as a part of the experimental procedure (see Appendix A). To standardize the pain induction within the experiment, refrigerated circulators [1] as special equipment from sensor technology, were used. These circulators cool the water at exact the defined temperature and constantly move the water within the bowl to prevent any warming of it in the area close to the hand. Varying the exact water temperature and the duration of immersion allows simulating different health states, concerning pain intensity and pain duration. With the cold pressor test, the induction of pain similar

---

[1] Julabo F12-ED Refrigerated/Heating Circulator;
http://www.julabo.de/us/p_datasheet.asp?Produkt=F12-ED 15.6.2013

to chronic pain is possible, a health state with significantly reduced quality of life. Accordingly, we generalize from the induced pain sensation and assume - that for the time of the experiment - the level of well-being the subjects experience can be controlled and varied with the cold pressor test. We can shortly sum up that with the cold pressor test a methodology was chosen which is easily applicable, reliable on the basis of the sensitive coolers, ethically acceptable, easily replicable and highly correlated with relevant health states.

Monetary consequences

Using the cold pressor test, the individual health state becomes a parameter that can be manipulated individually within the experiment. Within the conducted studies, health-related consequences were combined with monetary consequences. The experimenter asked for the individual willingness to pay to avoid the pain or to choose between different pain induction scenarios and to pay for the preferred one. Concerning the willingness to pay for different options, the design ensured that indicated amounts of money were realistic as subjects knew that in some cases, this money was subtracted from the payment they received for their participation in the experiment. The exact procedure is explained in more detail in chapter 1.

By means of the three experiments, new findings can be presented in the complex context of medical decision making. This includes the behavior of care givers on the one hand and preferences of the needy on the other hand. Implementing the cold pressor test is a new approach in health economics to enlarge findings from other studies working with hypothetical settings. By means of studies with differing designs and methods, the understanding of affected individuals and interactions can be increased and decision making in the health care sector can be built on a more sustainable fundament.

Chapter 1

1. Does your physician see your needs: an experimental analysis on prioritizing patients

1.1 Motivation: scarcity of resources in health care

In Western countries people are used to health care at the current state of the art accessible to every needy individual. Especially where statutory health insurance is established, patients are often not aware of the costs that result from their treatment {Hauerwaas, 2008}. At the same time we can observe rising demand for a health care which maximizes well-being far beyond medical care to ensure survival. At least partly services such as dental care, sports courses or artificial insemination are refunded by health insurances. At the same time, the higher life expectancy results in an intensified purchase of health services of a growing number of elderly. Also the number of multimorbid very old patients is rising (Scholl et al., 2013). Expensive treatments after the age of 75, for example heart surgeries or new hip joints followed by enlarged rehabilitation durations represent a rising expense factor.

As a consequence, politicians, care givers and increasingly society as a whole discuss the question how we can accommodate the growing demand for health care services with the limited resources from statutory health insurances. The first step is the optimal usage of available resources and the identification of potential for optimization. However, this so called rationalization (Wernitz & Pelz, 2011) to reach the targeted efficiency aims is not enough. It is not any longer possible to deliver every health care service to everyone at any time (Schultheiss, 2004). Accordingly, the sequencing step after rationalization is prioritization. Prioritization means the definition of disease clusters or patient groups to which resources are allocated first. Prioritization must strongly be differentiated from rationing where we do not ask who gets the treatment first but who gets the treatment anyway (Fuchs, 2010).

What remains problematic in terms of prioritization is its definition for a smooth realization within daily routines in hospitals or medical practices. Decision makers have to ensure that they deliver the same treatments for comparable symptoms. Moreover, decisions must be transparent so that the affected individual can comprehend and accept it.

To show an example for early engagement in terms of prioritization, the following section focuses on the development in the state of Oregon in the United States. Oregon first made efforts in the formulation of a framework for prioritization in 1990 (Brown, 1991). Differing from approaches of other countries the years after, Oregon aimed at defining a core bundle of services available within health insurance designed for poor workers. As Oregon was an outrider aiming for the definition of core services, they faced big methodological difficulties. In the first place, a cost-benefit approach was implemented to come to a ranking of most necessary services. However, a disputable list was the result and as a consequence, the final definition of core services strongly based on judgments of the engaged commission (Ham, 1997) instead of the affected individuals.

After the state of Oregon, several other countries all over the world started priority setting approaches. In general, these efforts can be separated into two different branches referring to their target: Oregon, New Zealand, Israel and the UK wanted to come to concrete allocation decisions whereas Norway, the Netherlands, Sweden and Denmark started priority setting approaches aiming to outline abstract principles for prioritization (Sabik & Lie, 2008). Accordingly, the former group of countries came to concrete recommended services and effectiveness, efficiency as well as evidence postulations. Sweden and Denmark avoided concrete recommendations but compounded with general statements on human dignity, solidarity or equal human worth. Norway and the Netherlands came to questions on necessity, severity or effectiveness. Thus, the different results of the eight countries differ in abstract or concrete character as well as in their applicability.

The question that results from the efforts the above mentioned countries made to come to a good prioritization guideline is how the priority setting approaches can be evaluated and how they influenced political decision making so far. Sabik and Lie (2008) propose three aspects for the evaluation of the individual approaches, namely the inclusion of the public within the process of priority definition (1), the formulation of applicable principles in the prioritization decision including the cost factor (2) and the effect on policy and practice (3). The first point is mainly fulfilled, with UK for example implementing the Citizens Councils, the public meetings organized by the New Zealand Committee to discuss documents or the incorporated feedback from public surveys in Sweden and the Netherlands. More questionable is the inclusion of costs within prioritization principles (2). Here, only New Zealand emphasizes cost-effectiveness as a primary consideration. The other countries hesitate to clearly include the implementation of economic studies, as for example the appropriate measure for utility is not clear (Denmark). Israel postulates that costs should only be considered if all else were equal. The common sense is the maxim that access to basic health care for everybody is preferred over access to every treatment option for some (Marckmann, 2009; Zentrale Ethikkommission, 2007). The most important point is related to the consequences of the efforts the different commissions and councils had (3). The individual goals defined by every country were failed as no concrete applicable guideline could be developed. This would have been really helpful for everyday processes. The processes to define useful ethical priorities still continues. Guidelines remain abstract and broad so that difficult cases cannot be handled based upon them. Questionable services were not excluded from funding and political issues or media pressure indirectly influence decisions. Marckmann (2009) suggests that every country should implement stable institutions to head and coordinate the ongoing process realized in close cooperation with medical profession and science. This process must be legitimized by the public fostering constant communication, as the author explains.

The overview of these eight countries demonstrates how difficult the approaches have been so far to come to an applicable guideline for prioritization practices.

Besides, values and principles vary within different societies and cultural environments and the willingness to confront oneself with the resource dilemma in the health care sector is differing. In Germany for example, the question of transparent explicit prioritization procedures is slowly emerging during recent years and a new institution (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen) has been implemented to investigate evidence of medications and treatments to deduce recommendations for their funding. However, the German public is rarely integrated in resource allocation decisions (Marckmann, 2009; Zentrale Ethikkommission, 2007) and forums analogous to Norway or Denmark are still missing. In the light of scarce resources the necessity of prioritizing is rising, especially if we want to avoid an expanded exclusion of essential health services. Accordingly, in Germany we have to foster the dialog with the public and conduct appropriate studies to come to accepted prioritization criteria. Thus, we can preserve a health system still able to provide all necessary services for everyone.

1.2 Aspects of prioritization

So far, no country has defined applicable guidelines for a transparent prioritization and all efforts resulted in abstract formulations rarely considered in political decision making. The following sections present the actors in prioritizing in more detail. Subsequently, a focus lies on the influencing criteria for prioritization decisions and approaches how these decisions can be evaluated. Thereby, the status in Germany is analyzed in detail. In Germany, the situation deviates in some central factors such as the realization of public debates as will be shown in the following chapters.

1.2.1 Who prioritizes?

Depending on the individual health care system a country has established, there are different players with differing options to influence resource allocation and prioritization. In general, separable are macro, meso and micro level, i. e. national, local or individual decisions (Klein, 1993; Ham, 1997). The macro level comprises governmental decisions for example on how much of the national budget should be spent in the health care sector. It can also include international or regional resource allocation, for example in terms of higher financial contribution for specific regions or populations. The meso level includes a heterogeneous group of actors, such as health service provider agencies or health insurances. Decisions relevant for specific patient groups or individuals are summarized on the micro level. Mostly physicians decide on access to treatment, urgency for an intervention or selection of the adequate therapy (Ardal et al., 2008). In general, it can be differentiated if decision makers on the different levels, especially in meso and micro level, prioritize explicitly or implicitly. Whereas for instances the official exclusion of services or patient groups is a transparent and acceptable procedure for the affected individual, implicit prioritization means i. a. the delay, dilution or deterrence of services (Parker, 1975). Here, the affected patient often does not realize that he is facing the consequences of prioritization or rationing, and if so, he has rarely an option to influence the situation. This is the case when restricted budgets for a quarter are exhausted and no more prescription is written or one has to wait four weeks for the next available appointment to see the familiar physician.

In Germany, implicit and explicit prioritization over the different levels is splitted into nine different bodies (Wodarg, 2010). First of all, legislation defines cooperation between subordinate institutions as well as claims for health service. Government, by means of different institutes, realizes serious prioritization decisions concerning vaccination recommendations, provided medications, prevention targets or the funding of selected research projects. In 2004, Germany established the so-called *Gemeinsamen Bundesausschuss* (G-BA), a panel of four bodies (*Kassenärztliche* and *Kassenzahnärztliche Bundesvereinigung*, *Deutsche Krankenhausgesellschaft* and *GKV-Spitzenverband)*. Together, this panel organizes the realization of legal demands. Their decisions are binding for both, the insured individual and the care giver. The G-BA establishes the catalogue of health services refunded by the health insurances and thereby also makes the decision which treatments cannot be provided. Therefore, in 2004 the *Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG)* was founded to organize relevant research on new therapies and treatment success. The IQWIG receives order for necessary research from the G-BA. It is further allowed to work on own projects so that its influence on prioritizing comes from the selection of the specific projects and their research results. Another player in the health care sector in Germany is the statutory health insurance (besides the privat health insurance in which only a minority of the German population is insured) answering for the largest part of prioritization decisions in Germany. About 170 insurances compete in the German market and their competition is the explicit political aim to reduce costs. Not every illness is equally costly, huge expenses often attend chronic diseases for example. Thus, the statutory health insurances design contracts to attract or discourage the different patient populations. Moreover, specific patient groups are attracted for whom the insurance receives equalization payments so that it generates the maximal receipts. Accordingly, the umbrella organization of the statutory health insurances (GKV-Spitzenverband) further represents these interests. Moreover, associations and corporations have own interests on the health care market and influence the allocation of resources. Prioritization decisions comprise the closing of clinic units with reduced occupancy or arrangements between physician associations and health insurances concerning the exact diagnosis and the billing for the referring treatment.

One of the most evident players on the micro level are the individual care givers, especially the physicians. They organize themselves in the *Association of Statutory Health Insurance Physicians,* a body to represent the interests of physicians toward the government and the health insurances. Although ethic considerations exist, physicians can decide individually how to influence the organ allocation process or whether they offer home visits.

Besides the summarized instances, prioritization is also controlled by research, either at universities or the pharmaceutical industry: if no research is initiated to find treatments for certain diseases, appropriate therapies or medications will never get access to the catalogue of health insurances.

1.2.2 Criteria for prioritization

International approaches on the maxims for prioritization all conclude that first and foremost stands the ensuring of access to essential health care to everybody (Sabik & Lie, 2008; Zentrale Ethikkommission, 2007). To guarantee this availability of services, a society must desist from access to every possible treatment (Marckmann, 2009). However, the postulation to offer essential health care still excludes how this bundle of services can be defined. Efforts of countries highly engaged in a transparent prioritization (excluding Germany still avoiding an explicit debate) have come to two different frameworks to answer this question: Norway, the Netherlands, Sweden and Denmark were outlining principles (1) whereas Israel, New Zealand, Oregon in the US and UK were defining practices (2) (Sabik & Lie, 2008). The first framework was based on discussions with health care experts initiated by government. Some countries additionally included the public or government officials. The efforts resulted in guidelines and recommendations for priority setting with a strong focus on ethics and values. Examples are criteria such as human dignity and solidarity set in Sweden or equal human worth and freedom defined as criteria by Denmark. With a focus on practices, the second approach is shaped by more concrete applicability of criteria. Examples are effectiveness, efficiency, equity and acceptability summarized as aims by New Zealand or Israel presenting a hierarchy of criteria starting with Life-saving technologies with full recovery until the funding of efficacious treatment that is expensive to the individual but of reasonable cost to society. Although the second approach is less abstract, the influence of both frameworks on politics and prioritization decisions is rare. The mentioned countries have increased public awareness but did not manage to implement the desirable standardized procedures (Sabik & Lie, 2008).

On the way to prioritization criteria, a central instrument which is used in almost every engaged country is the involvement of the public with information campaigns or discussions (Buxton & Chambers, 2011). In Great Britain, the NICE initiated one of the most promising approaches with the installation of a standing body, the Citizens Council, formed of 30 individuals without medical background. This council represents the general population in terms of age, socioeconomic background, gender and ethnicity (NICE, 2008). The NICE intends to include the opinion of the public within its decision making in terms of treatment, prevention and health promotion. The aim of approaches involving the general public inside and outside Great Britain is to come to useful criteria for decision on need and fairness and also to generate acceptance for the consequences of limited resources. Although the process of public involvement is valuable and desirable for countries like Germany (Marckmann, 2009) there is evidence that allocation propositions presented by representative samples of the public do not come to fair and efficient solutions. Ratcliffe et al. (2005) investigated the efficiency of liver allocation decisions of the public. They find that equity concerns significantly direct their decision. As a result, the efficiency determined with QALYs is reduced and also patients with low survival probabilities receive an organ. Green (2009) presents a study on the relevance of the criterion *severity of health* for public allocation preferences. Analog to Ratcliffe et al. (2005) he comes to the result that equality is a strong allocation tendency. Subjects prefer an equal distribution between patient groups of different health state severity although the potential health gain for these groups varies significantly. Subjects demonstrate at least equal allocation preferences for the group that is more severely affected. This means that a unit of health gain is of greater social value given to a disadvantaged patient group. A representative study on the British population including 559 individuals dedicated to the relative societal values of health gains again comes to preferences deviating from QALY-based solutions (Dolan et al., 2008). Related to this report, six general findings can be summarized:

- First, comparable to Ratcliffe et al. (2005), individuals demonstrate a general inequality aversion and thus also support resource allocation to patient groups with reduced chances to profit.
- Second, requested individuals evaluate the timing of illness to be decisive and claim that children should be prioritized.
- Third, there is a preference for those patients severely ill so that they should receive treatment first.
- Fourthly, individuals want the causes of illness to be taken into account so that own responsibility in comparison to bad luck should be posteriorized.
- Fifthly, an influential factor for resource allocation is the labeling of conditions: here for example illness in the context of obesity is posteriorized.

- Finally, subjects indicate a preference for patients suffering from extremely rare conditions.

Besides the approaches to prioritization criteria on a national level including public allocation preferences, in the next step the concrete allocation behavior of care givers is presented. Here different factors are taken into account within relevant studies. However, the picture the studies draw is heterogeneous. Factors are for example age, socio-economic information, health-related parameters and treatment costs. While van Delden et al. (2004) demonstrate that medical decision makers reject resource allocation based on age, others report the general favoring of children or the posteriorizing of demented or attended elderly (Strech* et al., 2008a; Ryynänen et al., 2000). In terms of the socio-economic status of the patient, there is evidence that physicians prioritize patients with a higher social contribution (Strech* et al., 2008a) or that they accept social criteria under certain circumstances besides medical parameters (van Delden, 2004). Concerning the disease and the treatment, the related literature comes to the result that there are various details the care givers take into account, i. e. the compliance of the patient (Strech* et al., 2008a), the origin of the disease or its severity and prognosis (Ryynänen et al., 2000). Strech* et al. (2008a) as well as Ryynänen et al. (2000) also report that the costs of the treatment and thereby its cost-effectiveness play a decisive role for their resource allocation. In their review, Strech, Synofzik and Marckmann (2008b) summarize the themes essential for prioritization and cluster them in three categories: context-related (1), doctor-related (2) and patient-related (3). Besides other parameters, they report that physicians have to take into account the operating budget as well as the availability of resources (1). The doctor herself feels reluctant to withhold services and can hardly maintain a consistent standard of care while she aims to provide every patient with the best treatment (2). Concerning the patient the authors mention a demanding mentality as well as their ability to exercise pressure as relevant for allocation decisions. Additionally, preferences and personal circumstances of the patient can be decisive (3).

1.2.3 Evaluation of allocative efficiency: Quality-adjusted-life-years and Willingness-to-pay

In health economics and health policy, decisions must be made which are of massive influence for affected individuals. Especially patients are affected who suffer from a bad health state. Thus, it is of high importance not only to comprehensively shape the framework of allocation, but also to evaluate and adjust implemented procedures. This evaluation can focus different levels, be it the allocation of budget for research on a specific disease or the spending of money to finance a new magnetic resonance scanner in a local hospital. In this context the question is which allocation decision generates the best result for the majority of patients. The central problem for this specific economic evaluation is the definition and measurement of *the best*. What can be comprised here is life time but also well-being and quality of life respectively, with the latter treatment aim not necessarily reconcilable with the aim to deliver a maximal extended life time. Accordingly, there is not one analysis used to determine allocative efficiency, but at least two, each with its individual methodology: Cost-Utility-Analysis (CUA) and Cost-Benefit-Analysis (CBA) (Mitton & Donaldson, 2003). Cost-Utility-Analysis refers to the benefit from a program to standardized utility units. Utility is mostly defined by means of the QALY-concept (quality-adjusted-life-years) (Weinstein & Stason, 1977)[2]. The QALY-concept is based on the idea that every change in quality of life can be measured in gained or lost life years lived in perfect health. Thereby, different treatment results and effects of diseases can be compared on the basis of a similar unit of measurement, i. e. years in perfect health. With QALYs, the evaluation of different prioritization decisions is possible. Here, for example the health gains of different patient groups are contrasted with the necessary monetary investments. A possible result can be that for 1.000.000 Euros, a greater health improvement can be realized for 100 patients with illness X than for 100 patients suffering from illness Y. Thus, if treatment for illness X is prioritized, the achieved utility is higher. What remains problematic in Cost-Utility-Analysis is the definition of utilities. Different procedures can be implemented to derive QALYs. One option is the implementation of the abstract and time-costly operations time-trade-off or standard-gamble. Both options are conform to the axioms of the underlying von-Neumann-Morgenstern-

[2] Please notice that further details about the QALY concept will be presented in detail in chapter 3.1

utility-theory (von Neumann & Morgenstern, 1944). An alternative are questionnaire-like inventories which can be disease-specific or generic. Although their manageability is much better and even group settings are possible, their conformity with utility theory is controversial and the QALY-indices generated with different inventories vary significantly (Conner-Spady & Suarez-Almazor, 2003; O'Brien et al., 2003). Depending on the specific reason for an economic evaluation of the utility of a program there is a more general problem of the QALY-concept. The results are always utility indications referring to a concrete setting, patient groups, diseases or at least a field within the health care sector. Thus, Cost-Utility-Analysis suffers from a restricted comparability in terms of budget allocation decisions on the macro level. This means that for example it gives no information to answer questions whether people perceive greater value in a program increasing quality of life for patients, spending money to enlarge safety in traffic or to protect environment. In such a context, different units are necessary as QALYs do not offer any option for a sector-overlapping comparison.

With Cost-Benefit-Analysis, specific evaluations - even cross-sectoral - can be conducted because measurement of outcomes and values is realized in monetary units (Cookson, 2003). Coming from decision making in policy, Cost-Benefit-Analysis was first implemented in the 1960s and was continuously used to determine the benefit of measures in environmental (Davis, 1963) or traffic-related spending (Jones-Lee, 1974). Acton (1976) was the first to implement this analysis in the health care context. Besides the advantage to answer sector-overlapping questions of resource allocation, Cost-Benefit-Analysis is not based on utility-theory and does not need complex instruments such as time-trade-off or standard gamble to determine utility. Mostly, Cost-Benefit-Analyses in the health care context are implemented with contingent valuation to determine either the individual willingness to pay (WTP) to avoid or maintain a certain health state or the willingness to accept (WTA), the sum the individual needs to receive to tolerate a certain health state (Drummond et al., 2005). Although the QALY concept is more often used in health-related research, during the last two decades, there was a notable amount of publications using willingness to pay in health-related questions (compare for example Unützer et al., 2003; Chuck et al., 2009). The central advantage of Cost-Benefit-Analysis with willingness to pay in comparison to Cost-

Utility-Analysis is the possibility to make evaluations in a complex world where not all alternatives or informative details can be included. This is the everyday situation of all actors engaged in health care, where no aspired outcome occurs for sure and complications or side effects must be expected. To realize a Cost-Utility-Analysis and thereby maximize utility, complete information on competing programs is relevant. Moreover, all possible programs must be assessed at the same time. More common is the discussion of a few programs and their comparison within a concrete situation. Therefore, Cost-Benefit-Analysis is appropriate. Here, each option is represented by a monetary value and comparability also within a restricted area can be achieved (Drummond et al., 2005). Comparing different treatments by means of willingness to pay, patient’s preferences can be standardized and a clear preference hierarchy evolves. The interpretation of the absolute values subjects indicate as their willingness to pay is controversial and can result in unrealistic predictions (Cookson, 2003). However, for comparisons of different options, be it general evaluation questions or complex resource allocation problems, willingness to pay is a useful means and is worth further methodological investigation (Donaldson & Shackley, 2003).

1.3 Motivation and essential aspects of the experiment

Prioritization decisions on the micro level can hardly be evaluated for fairness or efficiency, especially when it comes to implicit prioritization. Here the patient is often not aware that services are reduced or postponed because of budget restrictions. A deeper understanding of the criteria that influence a physician during his bedside allocation can be investigated by means of an experiment including both, physicians and those in need for health care. In the following, an experiment is described investigating prioritization behavior of physicians over a group of different individuals in need.

Individual attributes of the needy

Within the conducted experiment, physicians had a budget at their disposal too small to help all presented subjects in need. As a basis for their prioritization decision, eight attributes were requested from the needy and presented to the physician. These parameters were related to the individual level of reduced well-being. Therefore, subjects were allocated to differing doses of pain induced with the cold pressor test[3]. Further parameters for the prioritization decision were the socioeconomic background and the willingness to pay to avoid bearing the imminent bad well-being under cold pressor pain. Although it is relevant to generally find out which of these eight attributes is recognized more to come to a prioritization decision, the special interest of this study lies on the usage of the information on individual willingness to pay. Besides the objective information the physician has about the different health states (pain intensity, duration of pain), with willingness to pay he has a subjective information about the perceived level of sufferance reflected in an amount in Euro. As Cookson (2003) points out, the strength of the method willingness to pay lies in the good comparability of monetary units: the physician can see the different indicated amounts subjects are willing to pay and he can deviate a hierarchy of urgency for treatment (Donaldson & Shackley, 2003). As a result, there is one hierarchy of treatment urgency based on objective health-related data (pain intensity and pain duration represented in water temperature and immersion duration) and a second hierarchy based on subjective urgency information deviated from amounts subjects indicate to pay for pain avoidance.

Physician's usage of information on individual willingness to pay

Information about individual willingness to pay is unusual to physicians in a concrete situation where they have to prioritize real patients. Nevertheless, the condensed information represented by an individually defined amount of money can have a mentionable supporting character to the process of decision making. The question is whether medical decision makers can capture the inherent information, if they are able to integrate it within the other details they know about the individual and if they are willing to work with information that is representing subjective perception. An investigation presented by Schattner et al. (2004) elicits that for patients it is an important question if the perception and concrete situation of the individual is realized by the physician and whether this is also influential in terms of

[3] compare the Introduction for further details about the pain induction method *cold pressor test*

the treatment process. Schattner et al. (2004) demonstrate that there are general differences in what physicians believe to be relevant for the patient and what the patient actually indicates to be relevant for him. Most striking is the aspect of patient's autonomy and physician's attentiveness to patient's preferences and rights, which is essential for the patients but not for the physicians. Autonomy contains the demand that information from the patient as well as her decisions in the context of the treatment must be taken seriously. Consequently, the physician should be aware of the individual situation, the perceived urgency for moves or the trade-offs the patient is willing to make to ameliorate certain aspects of his health. A means to operationalize this complex construct is the elicitation of the patient's individual willingness to pay.

Methodological challenges

The first challenge refers to the request situation of individual willingness to pay. What is specific for the conducted experiment is twofold: first, physicians knew the individual willingness to pay of the needy besides socioeconomic and health-related variables. Second, physicians knew the instant consequences of their prioritization decision for all involved individuals, either pain or no pain. Subjects in need knew that the willingness to pay they indicated could actually reduce the earnings they receive for participation in the experiment. Thereby the design contained a controlling condition to make sure that no high amounts were indicated just to influence the decision maker[4]. A number of studies demonstrated that a hypothetical character of the willingness to pay request influences the results. Blumenschein et al. (2001), for example, conducted a field experiment with asthma patients who could either buy an asthma management program for real or just hypothetically. The authors summarized that amounts of willingness to pay subjects indicated are higher in the hypothetical condition than in the real one. Analogous are results from an experiment investigating real and hypothetical willingness to pay and to accept repeated pain induction with the cold pressor test (Pesheva et al., 2011).

[4] For a detailed explanation of the procedure please compare chapter 1.5

A second challenge exists in terms of the format of studies concerning the implemented method to understand decision processes or view points. The work context physicians face in many countries contains a dilemma. In fact, three perspectives must be optimized. These are demands of the health care system for example in terms of budget constraints, own economic interests of the physician himself and the interests of the patients to receive the best care available (Strech et al., 2008b). More than once, these perspectives can hardly be combined. Physicians are discontented with the situation and face criticism when their economic motives are recognized. So it is questionable if their real decision behavior can be elicited by means of questionnaires or interviews. The answers they give when being directly asked might not reflect their real prioritization behavior. As a consequence, results elicited in studies excluding decision making tasks with real consequences might be confounded. This also includes the before mentioned studies which all work with structured interviews, discrete choice experiments, conjoint analysis or qualitative random paired scenarios (Ryynänen et al., 1999; Ryynänen et al., 2000; Schwappach, 2003; van Delden, 2004; Diederich et al., 2012). In these studies, deciders retrospectively report about their general behavior, reflect about hypothetical cases or try to imagine how they behaved in the role of a physician.

Summarized approaches rely on the self-perception of the individual or even on the willingness to reflect on and indicate own decision motives, which cannot be presumed. In the context of age-related prioritization, for example, the questioned physician might be afraid to face criticism if she confirms to use information on age to prioritize. It could be categorized as an unethical behavior to discriminate the elderly or to favor the working population. Especially in an interview situation, a social desirability bias can influence the individual answers so that physicians palliate the actual routines at the bedside (Adams et al., 1999). Moreover, it must be taken into account that because of the described dilemma the physicians are confronted with, there might be recurrent situations in which implicit values play a decisive role. The physician might not be aware that available information such as socioeconomic background or gender has an implicit influence on her behavior so that she slightly favors the poor or demonstrates a preference for educated patients, just to pinpoint two examples. Van de Martel (2008) asks for further research with additional methods such as experiments to come to a better understanding of prioritization procedures. As a consequence, the conducted experiment implements two methods to directly observe the decision process of the physician aiming to control for the two influencing mechanisms, i. e. a social desirability bias and implicit prejudices or unconscious decision schemes. These two methods are a self assessment using a ranking over the decisive relevance of patient attributes and a software analogous to eye-tracking observing the allocation of attention during the decision making process.

1.4 Research questions

Within the studies on criteria for prioritization presented beforehand, there is evidence that the different attributes of patients strongly influence the allocation behavior of the physician. This can be a demanding behavior creating urgency as well as gender or the family situation. Within the conducted experiment, the decision maker had eight different attributes of each individual in need at her disposal to come to his prioritization decision. There were health-related attributes randomly allocated by the experimenter at the beginning of the experiment (pain intensity (1) and pain duration (2) induced with the cold pressor test), the costs (3) the prioritization of each individual in need creates (also randomly allocated to the individuals), the willingness to pay (4) of the needy to avoid the assigned pain doses, an indication of each person if they were smokers (5) or not as well as general information in which city they were born (6), if their parents were academic (at least one parent) (7) and gender (8). For various reasons, this bundle of information might cause prejudice or preferences for certain individuals, be it sex or local origin. On the other hand, there might also be a strict focus on objective attributes such as pain intensity from water temperature. Accordingly, the research question here focuses on the usage of the different attributes for the allocation decision.

Individual attributes of patients:

1. **Which attributes of a group of individuals in need does a resource allocator use to make her decision?**

In this study, with willingness to pay a standard methodology from Cost-Benefit-Analysis was implemented. This parameter is related to money and thus might be perceived similar to the costs by the decision makers. Moreover, it includes information about the individual urgency not to receive the imminent reduction of well-being under cold pressor pain. Thus, a special focus lay on the usage of the parameter *willingness to pay* versus *costs.*

Willingness to pay:

2. **In comparison to the individual costs every needy generates if treated, what role does their willingness to pay to maintain a good level of well-being play for the allocator in her decision making?**

Relevant literature contrasting attributes of physicians and non-physicians find evidence that decisions significantly vary between these groups. This includes the evaluation of how important aspects such as the autonomy of the patient are (Schattner et al., 2004). Reyna and Lloyd (2006) elicited that more professional physicians process fewer dimensions of information and make sharper all-or-none decisions than less experienced colleagues. As a consequence from such findings, in the conducted experiment, two different groups of deciders were requested to allocate the budget to the group in need: not only prospective physicians but also students without any medical background. As explained before, individual willingness to pay amounts include information on the perceived urgency to get the questioned good. Especially within a standardized experimental situation where several individuals in need are in the same condition and indicate their willingness to pay, a high level of comparability is ensured. In our experiment, we provided the participants the possibility to indicate the maximum amount of willingness to pay. This maximum amount was restricted and was related to the prospective earnings for the participation in the experiment. As a consequence, limitations of the methodology willingness to pay are controlled, especially concerning varying individual financial situations (Donaldson & Shackley, 2003).

In terms of different allocators (medical and non-medical), the following research question was investigated.

Decider has / has not a medical background:

3. **Is the usage of willingness to pay information implemented equally for the prioritization decision**
   (1) if the allocator is a prospective physician or
   (2) if he has no medical background at all?

Within the broader context of scarce resources in the health care sector, considerations of efficiency in terms of resource allocation represent a central component. However, on the micro level within the interaction of physicians and patients, this factor is often difficult to investigate. In the presented study, willingness to pay amounts are elicited for all potential care receivers. By means of this information, the finally made allocation decisions can be evaluated for efficiency aspects. Therefore, two theoretical sums from stated willingness to pay are compared:

a) Amount of theoretically gained money from willingness to pay the individuals in need (in following figure called *patients*) who finally received the treatment had indicated to pay

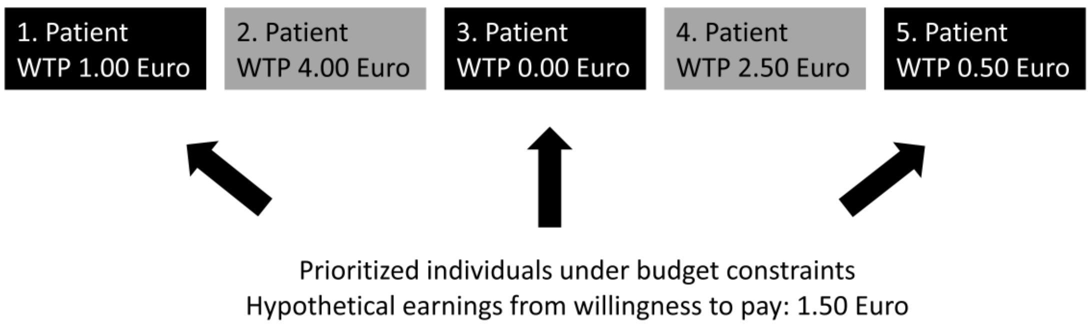


Figure 1 Possible willingness to pay distribution

Versus

b) Amount of theoretically gained money from willingness to pay if those individuals with the highest indicated willingness to pay would have been treated

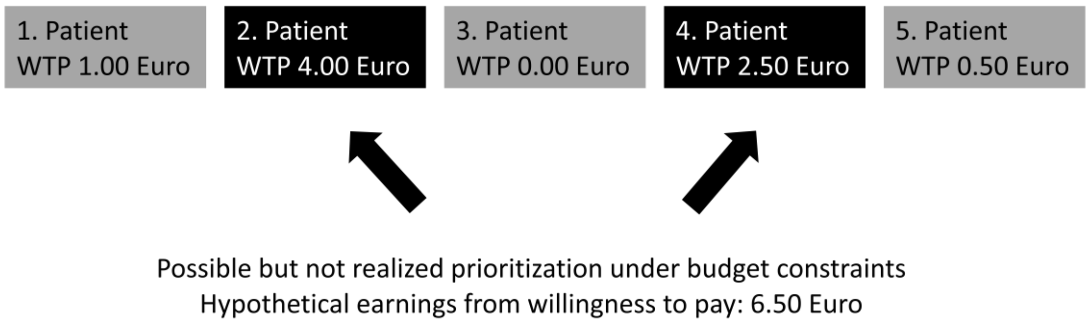


Figure 2 Alternative willingness to pay distribution

Analogous to the research question described before, the investigation of an efficient prioritization decision includes a comparison of the two decider groups that is prospective physicians and students without any medical background.

Efficiency

4. **Based on the individual willingness to pay, how efficient is the allocation solution the prospective physician makes compared with the decider without any medical background?**

1.5 Experimental design and procedure

The experimental design is based on the decision situation a physician is faced with in clinics: individuals in need for care vary in sex, origin and other personal attributes as well as in severity and endurance of suffering, costs of treatment and individual urgency. The experiment was conducted at the medical faculty of Otto-von-Guericke University in Magdeburg. The sample of subjects for the experiment consisted of two separable groups: the medical decision makers (16 individuals), and 176 patients – healthy students from all fields of study (subsequently called *the patients*). The medical decision makers were all medical students in the end of their studies who currently went through an interval of practical work in clinics. All of them had an overall work experience in the hospital of at least half a year.

At the time when subjects arrived in the lab, within half of the sessions altogether 16 individuals (one in every session) were randomly selected and were defined to be decision makers. These were the sessions in which the prioritization decisions were not made by real physicians but by subjects without any medical background. No subject was aware of the allocation procedure. In total the sample thus includes 16 medical and 16 non-medical decision makers and all in all 160 patients. The groups of all decision makers and patients were organized by the experimenters in way that they never met, neither before nor after the experiment to avoid any bias caused by real social interaction or potential interaction after the experiment. 32 sessions were conducted with one decision maker and 5 patients each. Decision makers and patients did not meet each other at any time during the whole experiment but were invited to different rooms on the campus. Before the experiment began, a subject number was assigned randomly to every patient using an urn with five balls. Thereby, specific levels of well-being could be allocated to every participant (water temperature and immersion duration with the cold pressor test), relevant and unchanged for the whole experiment.

The following combinations of temperatures and immersion durations were allocated. The specific combinations of time and temperature were selected in a way to make an easy intuitive hierarchy of the different levels of sufferance difficult.

- 12° Celsius for 3 minutes
- 9° Celsius for 2.5 minutes
- 7° Celsius for 2 minutes
- 4° Celsius for 3 minutes
- 4° Celsius for 1 minute

All subjects started the experiment at the same time. One experimenter observed the decision maker in the one room and the second experimenter coordinated the five patients in the other room. The experimenters distributed a detailed instruction equal for both groups. The usage of any words referring to a medical context was avoided so that the deciders were labeled as Type-A players and patients as Type-B players (compare Appendix A). Both groups were aware of the whole procedure right at the beginning of the experiment, for themselves as well as for the other group, decision maker or patient, respectively. It was explained to the patients that within the experiment their level of well-being would be reduced by pain induction in case that the Type-A player does not set them free using his limited budget. On their screen, the patients were informed about their individual level of reduced well-being awaiting them, water temperature and immersion duration. After signing a consent form for the participation in a pain experiment which all participants did, subjects sequentially tried a cold water bowl of 7° Celsius for 20 seconds. As the allocated water temperatures were 4°, 7°, 9° and 12° Celsius, 7° C was evaluated as a representative sensation for all relevant temperatures. This trial procedure was essential to give subjects an impression about the possible pain experience which they could use to evaluate how the cold water feels like for them. This impression formed the basis so that subjects could state their willingness to pay to avoid the experience of the allocated pain doses later in the experiment. The decision maker also went through this trial phase to make sure he could imagine how the pain feels alike.

Patients: Subsequently, the pain-related (assigned temperature and immersion duration) as well as demographic information (gender, birthplace, academic background of parents, smoker or non-smoker) was collected from every subject using a computer-based questionnaire. The experiment continued with the willingness to pay request to get to know how bad each individual evaluates the pain she might face later in case that the decision maker did not set her free. This request could end up in quite different amounts for strong pain levels and also for low pain levels, depending on the individual decision of each subject. For example a patient allocated to an intermediate immersion duration in warmer water (for example 9°C for 2.5 minutes) could indicate a willingness to pay of 5 Euros to avoid the treatment (the maximum possible willingness to pay), whereas a patient facing a longer immersion duration in colder water (for example 4°C for 3 minutes) could have indicated not to be willing to pay at all to avoid it. The willingness to pay request was designed as follows: column A contained the individual pain doses whereas column B contained rising amounts of money (0.00 – 0.50 – 1.00 – 1.50 - … - 5.00 Euros) (compare table 1).

	A: Receiving pain doses	B: Paying amount of money	Choice: A or B
1	1 minute in 4° C	0.00 Euro	
2		0.50 Euro	
3		1.00 Euro	
4		1.50 Euros	
5		2.00 Euros	
6		2.50 Euros	
7		3.00 Euros	
8		3.50 Euros	
9		4.00 Euros	
10		4.50 Euros	
11		5.00 Euros	

Table 1 Procedure to define maximum willingness to pay

Decision scheme to define maximal willingness to pay to avoid the pain. In the second column, the individual temperature and immersion duration were indicated for every patient. The decision maker was informed about the willingness to pay amounts individually determined to use this information for his allocation decision.

The patients indicated for every row, if they preferred to pay the money or to receive the pain doses, so that there was one switching point for every patient, the earliest rational switching point between a willingness to pay of 0.00 and 0.50 Euros. Subjects were informed that in the end of the experiment, one of them would be determined randomly and one of his eleven willingness to pay decisions would be chosen and realized. This means that depending on his decision, the chosen individual either had to go through the cold pressor test with his water temperature and duration or his payment for participation in the experiment would be reduced according to the stated amount he was willing to pay in the defined case. After the willingness to pay was indicated, the costs to free the individual from his pain doses were defined. To define the amount decision makers had to take from their fixed budget of 5 Euros (an experimental budget not related to the payment of subjects) to set a patient free, a fix algorithm was used. This algorithm allocates prices so that a maximum of three or four patients could be prioritized. There was no direct relation between willingness to pay, temperature and immersion duration and the allocated costs to free the individual from pain.

In the end of the information request phase, eight attributes were collected or allocated for each of the five patients:

1. Allocated pain intensity
2. Allocated pain duration
3. Allocated costs to treat this patient so that he will not face pain
4. Individually indicated willingness to pay on the basis of the allocated pain doses
5. Indication if the person was smoker
6. Information in which city the person was born
7. Information if their parents were academic (at least one parent)
8. Gender

Decision maker[5]: After awaiting the eight parameters from the patients, the decision maker could see on his screen a table containing these parameters for each of the five patients in the group. To objectively analyze which attributes of the patients are really relevant to come to a prioritization decision, simple requests are not enough but can be influenced by factors such as social desirability. Therefore, an instrument to analyze on which attribute subjects concentrated was implemented: the software Mouse Lab (Martijn Willemsen & Eric Johnson, 2006). During the decision phase, all attributes were covered with labels such as "sex" or "temperature" (compare figure 3). A touch with the cursor on the field of interest opened it and the information became visible as long as the cursor rested on the field. Thus, there was a data generation on how often which field was opened and how long subjects left it opened. The decision makers had as much time as they needed to go through the attributes again and again to decide finally whom to free from the allocated imminent pain experience. The restricted budget allowed freeing a maximum of three or four patients.

[5] *decision maker* always refers to both groups, prospective physicians and students without medical background

Figure 3 Mouse Lab surface

Surface the decision maker saw displayed with Mouse Lab. All attributes of the five patients were labeled and could be opened using the cursor. In the text below the labeled fields, the budget (5 Euros) was displayed and the decision maker had to calculate whom he could treat and for whom his budget would not last any more. After finishing the decision procedure, the decider could mark the numbers of the patients he prioritized on the screen.

After the decision maker had made her allocation decision, all subjects answered a questionnaire about the priorities a decision maker could use for her decision whom to free from pain. The questionnaire varied for both groups in framing only. It asked decision makers to rank the eight attributes of the patients (pain intensity (1) and pain duration (2), the costs the prioritization of each individual in need creates (3), the willingness to pay of the needy to avoid the assigned pain doses (4), an indication of each person if they were smokers or not (5), general information in which city they were born (6), if their parents were academic (at least one parent) (7) and gender (8)) on relevance for their decision, starting with the most important one. Deviating from that, it asked patients to speculate about the relevance of the parameters for a decision maker and to rank them accordingly. Thereby, a completely parallel experimental procedure could be realized both for patients and decision makers. However, the focus of the study lies on prioritization behavior of decision makers and as a consequence, the data from the patient questionnaire was not included in the further analysis of this experiment.

Patients: As soon as all patients had finished the questionnaire, the experimenter informed them who was prioritized and who would have to go through the pain experience. At that point, the experimenter in the laboratory of the patients started a video camera focusing on the water bowls. The camera transmitted the pain induction procedure to the laboratory where the decider sat. This procedure allowed for a live observation of the decision consequences the decider had just made. The demonstration of the today's newspaper and the time displayed on a mobile phone proved that the pain induction procedure really happened in this moment. All affected subjects sat down at their water bowl with their individual temperature. The immersion began at the same time for all subjects and each individual was informed, when the time for her was over.

Decision maker: The whole procedure of the patients sitting next to the water bowls had to be observed by the decision maker via web cam. After the transmission of the pain induction for the patients was finished, the decision maker was paid 15 Euros for her participation in the experiment and she left the campus.

Patients: After the pain induction the experimenter used an urn to define one individual for whom the willingness to pay decision made at the beginning of the experiment was realized. The individual decision matrix (compare table 1) of the selected participant was opened and one of the eleven decisions was determined randomly and realized. This either meant the individual pain experience (a potential repetition of the pain induction if he had not been prioritized by the decision maker) or the payment. If he had been willing to pay to avoid the pain, the subsequent payment for the participation in the experiment (also 15 Euros as for the decider) given to all patients was reduced for him accordingly. If for example the indicated willingness to pay at the defined decision (not necessarily the maximum willingness to pay) was 1.50 Euros, than the payment for participation was reduced so that the individual received 13.50 Euros instead of 15.00 Euros.

1.6 Results

In the first place, the conducted experiment was designed to shed light on the attributes decision makers use to prioritize patients. Therefore, 32 sessions with a decision maker and five patients each have been conducted. The prioritization decision could be based on eight different attributes either indicated by the patient himself or allocated to him. Findings from preceding studies elicited that variables influence the allocation decision which can hardly be evaluated as objective criteria. Our expectations for the conducted study were analogous to older stories. For example, we expected the prioritization for individuals coming from Eastern Germany as the experiment was conducted there. Another possible expectation referred to preferences for females. Several decision makers articulated a belief that women suffer more from the cold. However, the results are not representing any gender-related preferences.

First of all, the evaluation of data from the ranking is presented and compared with the data generated using the Mouse Lab software. Decision makers were asked to rank the importance of the eight attributes they had at their disposal for the allocation decision. Therefore they could use numbers from 1 (very unimportant) until 10 (very important). The following figure 4 presents the ranking of the attributes for all the 32 decision makers. What is obvious here is that the objective health-related criteria water temperature and immersion duration are most important followed by the costs to prioritize the individual based on the restricted budget. Less relevant was the information about the indicated willingness to pay, the more subjective attribute about the perceived suffering caused by the allocated pain doses. This already refers to the second research question about the usage of the attribute willingness to pay in comparison to the other monetary information, the costs to free from pain. According to the ranking, willingness to pay is significantly less important for the decision makers than the costs (Wilcoxon, 1%-level). An analog picture occurs after the analysis of the Mouse Lab data. Even in terms of the comparison of willingness to pay and costs, a significant difference is evident (Wilcoxon, 5%-level). Interestingly, the remaining information seem almost not to be relevant for the allocation decision, gender plays a minor role. Smoking habits, an informative attribute that could take effect as either a reason to punish or to reward is unimportant to the mean decider, as well as the birthplace and the level of education the parents have (compare figure 4).

So the first research question that asked which attributes of a group of individuals in need a resource allocator uses to make her decision can be answered now. The basis therefore is the ranking representing the self-indicated decision procedure and the further investigation of these self indications using Mouse Lab data. Of importance for the deciders are the objective criteria water temperature and immersion duration referring to the objective level of well-being the patient experiences. Moreover, the deciders mention the costs to prioritize the individual patient and thus proof their strong focus on the economic component of their decision.

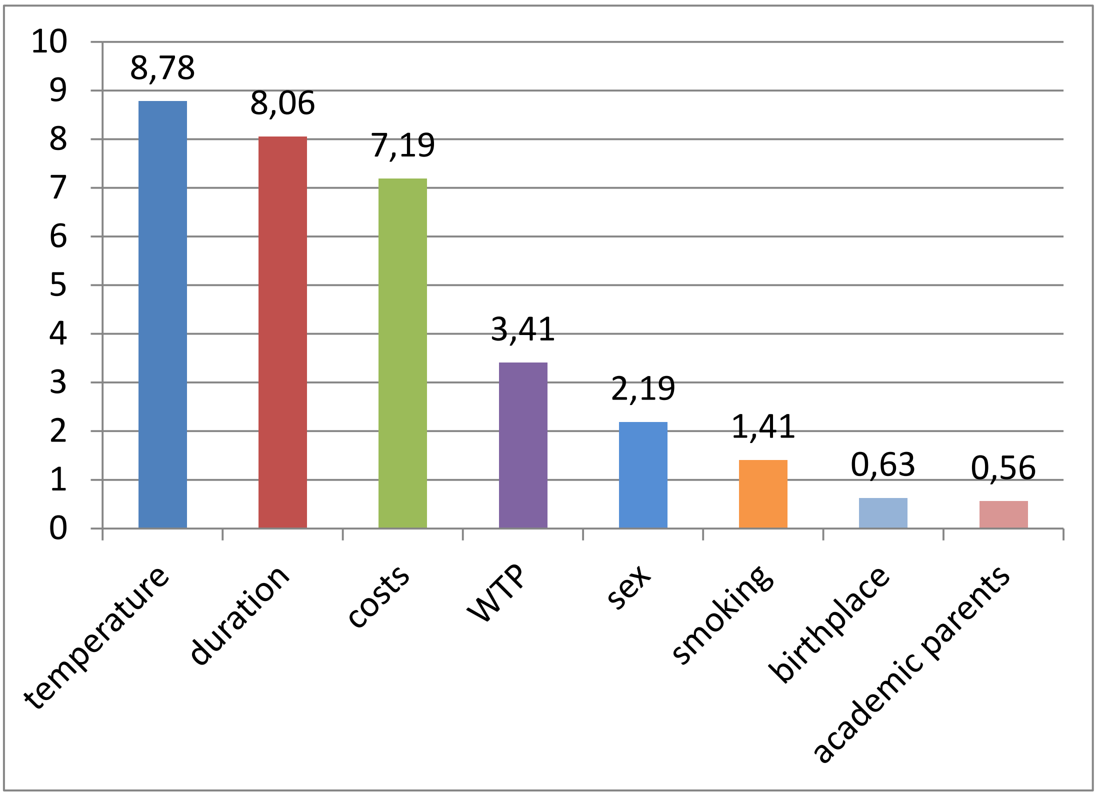


Figure 4 Ranking of Attributes

Results from the ranking the whole group of decision makers indicated to explain the importance of the available attributes to come to an allocation decision. 10 means high importance whereas 0 stands for no importance for the decision.

In the next step, the results from the separated analysis of prospective physicians and students of differing faculties are presented. Again, the focus lies on the explicit usage of willingness to pay according to the ranking. The individual amount patients indicated to pay to avoid the expected pain refers to the subjectively perceived decline in well-being the individual ties up the allocated cold pressor procedure. The assumption of a differing behavior of usage concerning willingness to pay seems appropriate for the two decider groups. Physicians go through a socialization where they are constantly confronted with an environment of need and sufferance. In fact, prospective physicians indicate to use willingness to pay information significantly sooner than deciders without medical school background (Mann-Whitney-U-Test, 5%-level, 1-sided) (compare figure 5).

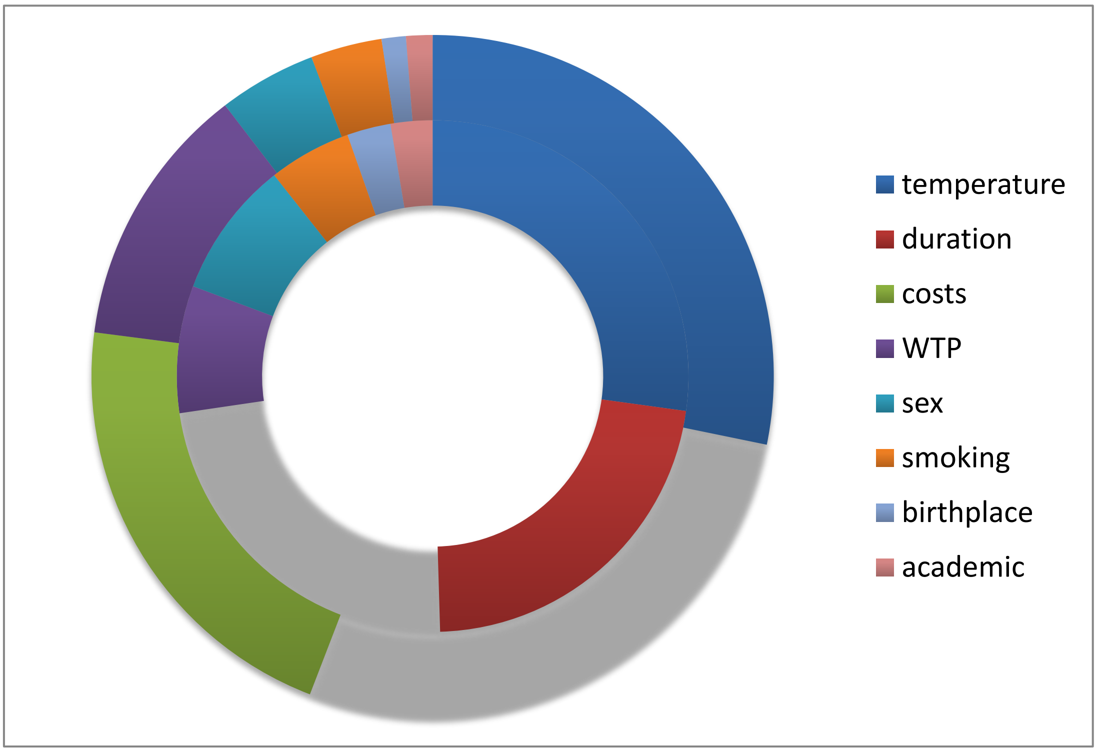


Figure 5 Attribute usage between decider groups

Comparison of usage of the available attributes to come to an allocation decision. The inner circle represents the group of deciders without medical background; the outer circle stands for the prospective physicians with working experience.

The last but most comprising topic to analyze within this experiment is the allocative efficiency. As for the usage of willingness to pay, the evaluation of achieved efficiency levels is done separately for prospective physicians and other students. The central idea in terms of efficiency is the theoretical earnings implicated with willingness to pay. To give evidence for the efficiency inherent in the allocation decisions, a comparison is conducted of these theoretical earning gained through the actually treated individuals with the theoretical earnings that could have been made if the individuals with the highest willingness to pay had been treated (compare figure 1 and 2). For the whole group of deciders, in 20 out of 32 sessions the efficiency had not been maximized. This creates an overall loss of efficiency of 25 percent. However, there is a relevant difference between the efficiency levels of the two groups of deciders. Non-medical students reach an efficiency level of 58 percent on average, whereas prospective physicians come to an efficiency of 93 percent (compare figure 6) and thus are significantly more efficient in their decision making (Mann-Whitney-U-Test, 1%-level). This means that the latter group maximizes utility for those in need much better than the non-medical students.

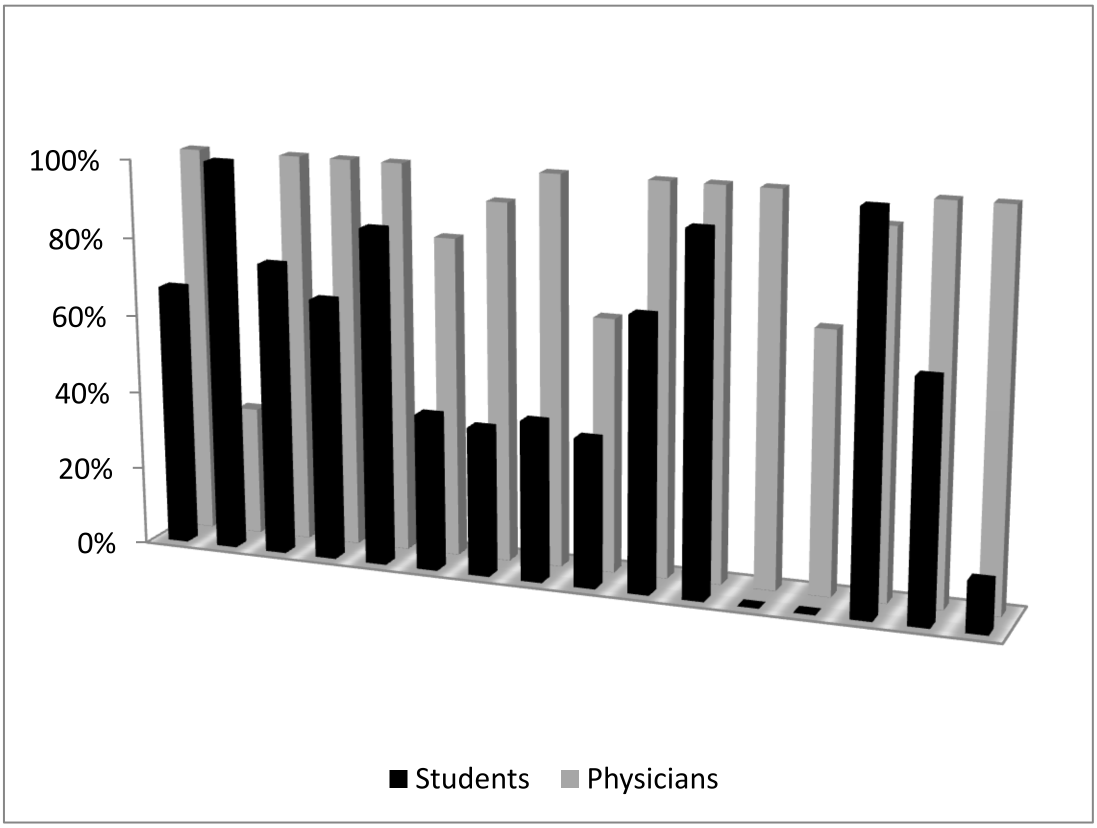


Figure 6 Comparison of efficiency levels

Every bar chart stands for one decider. The gray bars symbolize the physicians whereas the black bars represent deciders without medical background. The chart gives a general overview on the distribution of reached efficiency levels in both groups.

1.7 Conclusion

The presented experiment investigates the allocation process of physicians at the bedside. By means of an experiment which includes prospective physicians to make allocation decisions, the complex allocation scenario is simulated. Basis for the decisions were different attributes of the patients, namely personal, social, health-related and economic ones. The attribute of special interest within this study was willingness to pay and its usage for allocation decisions. Willingness to pay is a complex conglomerate of information including not only economic, but also aspects of individual sufferance and perceived urgency for treatment. Although its character in terms of a monetary amount suggests its closeness with the information of costs for treatment, both attributes are used significantly different. Referring to the ranking the deciders gave and based on the Mouse Lab data, costs are implemented in the decision making process more than willingness to pay.

The study demonstrates that attributes bound to discrimination such as gender or family background play a minor role within prioritization decisions. Moreover, allocators with medical expertise are more attentive to information such as willingness to pay than those without relevant experience. For all allocators, objective information of the health state as well as the costs to treat the specific individual is most important.

Concerning the investigation to understand the prioritization procedure every decider went through, two instruments were used: a ranking of the attributes for relevance and software analog to an eye-tracking instrument. Interestingly, these two instruments came to analog results concerning the attention dedicated to the different attributes. As the implementation of Mouse Lab is not necessarily easy to realize, it is important to recognize that also the standard instrument of a simple ranking comes to reliable results.

In terms of the efficiency of the allocation decisions, the experiment demonstrates that prospective physicians are already more efficient deciders than other students. They better focus the need of the patients represented in willingness to pay and thus maximize their utility better. This behavior can also be interpreted in the way that subjects with medical expertise have a higher trust in the evaluations of patients and are willing to respect their situation. A relevant finding from this study is the general acceptance of willingness to pay in the context of resource allocation. None of the deciders neglected to recognize the willingness to pay information.

Moreover, the aim of the conducted study was to simulate the high level of complexity a physician faces in her daily practice. Willingness to pay can deliver important insights into patient needs and also comprises factors not directly related to health. This can be an advantage in comparison to the implementation of QALYs. This is the case if the pool of potential care receivers is homogeneous. Additionally, willingness to pay can be a helpful method if maximal willingness to pay is defined proportionate to another entity like in this study the payment for the experiment. In such an environment, even the hypothetical character of willingness to pay can be reduced or completely switched into a real scenario.

Chapter 2[6]

2. Do people have a preference for increasing or decreasing pain?

An experimental comparison of hypothetical and monetary consequences

2.0 Motivation

The preceding chapter presented a complex experiment aiming to investigate a decision process in resource allocation. The focus is on the perspective of physicians prioritizing over a group of specified individuals in need. Different parameters concerning the situation of these individuals and their background formed the basis for decision making. The individuals themselves were not included in the process but had to accept their prioritization or posteriorization.

Budget restrictions are only one reason for decision making in the health care sector and in terms of care for individuals in need. Such decisions reflect scarcity. However, an increasing number of decisions must be made because of the richness of opportunities for different health situations. Advanced health care technologies or alternative treatment options as for example natural remedy offer new chances to adapt treatment plans to patients. On the one hand, the physician is still the instance to reflect over alternatives, his or her costs and the side effects the patient might have to tolerate. But on the other hand, new aims in patient care are to share relevant disease-related information with the patient, to communicate about the level of well-

[6] This chapter is to a great extent based on Kroll, Trarbach & Vogt, (2012).

being the patient experiences and finally to include her in the decision for the best treatment plan. The inclusion of patients in the decision process by means of his endowment with decision-related knowledge as well as the consideration of his situation and values is called shared decision-making (Barry & Edgman-Levitan, 2012). Besides additional aspects, shared decision-making means to explain choices in the course of treatment, to avoid or control for framing effects within the communication process over alternatives and to perceive and respect the role the patient wants to have within the decision process, including the ability and willingness to tolerate risks (Godolphin, 2009). Expected positive effects from shared decision-making are i. a. an increased health-related knowledge, a more adequate risk assessment, more decisions in line with the values of affected patients and as a result less internal decision conflicts for the patient (Stacey et al., 2011). Accordingly, communication between physician and patient becomes more challenging to the physician. However, a successfully realized communication supports the target to come to the best health state possible both within the treatment process and afterwards.

We know from various disciplines such as psychology and economics that decision making processes underlie biases which make individuals decide irrationally. As decisions in the health context often lead towards existential consequences for the patient, it is of special interest to research behavioral anomalies in this setting. Accordingly, different anomalies and biases are investigated and presented in this chapter.

The following experiment focuses on two aspects concerning investigations of preferences. The first aspect includes the choice situation patients often face when more than one treatment is considered. In general, treatments often can be differentiated in terms of their influence on quality of life during the treatment process. Cancer treatment for example can be either a surgery or a chemo therapy. The former often means a strong reduction of quality of life during the time of the intervention and the days after. A stepwise amelioration of well-being might follow as a standard development of the ordinary patient. The case of treatment with chemo-therapy can be the other way round. At the beginning, well-being is only slightly reduced, but it deteriorates more and more within the period of repeated medication. Hypothetically it can be assumed that in both scenarios, the overall level of well-being is equal and the only difference is the order of levels in quality of life until finally full health is re-established. Other diseases with possible treatments can also be differentiated by the distribution of levels of well-being over time. The general question is if patients have fundamental preferences about the timing of well-being during the process of reconvalescence.

Besides the often difficult decisions patients face in the context of their health and relevant treatments there is a second aspect investigated in this chapter. It refers to the methodology to research preferences individuals indicate. As described in the introduction of this work, real consequences from made decisions are one basis to reduce behavioral anomalies. In terms of patient decision making over real health situations, such an investigation is rarely ethical and remains either in a laboratory context (Pesheva et al., 2011) or refers to more irrelevant decision scenarios (Blumenschein et al., 2001). Nevertheless, such studies are important to add to the picture we have so far drawn for representing decision behavior and anomalies. Accordingly, the following study includes real consequences for participants, i. e. reductions of well-being induced with the cold pressor test and monetary disadvantages if preferred treatments are received.

Analogous to the experiment described in the first chapter, again in this experiment a wording free from any health-related content was chosen.

2.1 Introduction

To choose between two alternatives, a decision maker typically needs a preference for one alternative over the other. In case of two experiences, this choice is not simple: experiences rarely consist of one single event but they are combinations of different events or sequences of single outcomes. To derive the value of an experience, all events forming it have to be evaluated and the values of the events must be combined. For this complex procedure, different theories and predictions exist. However, it is not clear which theory captures human behavior best and corresponding evaluations with monetary incentives are – to the best of our knowledge – still lacking.

Let there be two meals for lunch: stale bread and steak. An improving (increasing) experience[7] would be eating stale bread on day one and steak on day two, while a worsening (decreasing) experience would be starting with steak on day one and eating bread the day after. For a rational risk-neutral decision maker, both experiences, the increasing and the decreasing, have the same overall value – the sum of the values of each meal. In other words, rational decision makers are indifferent over experiences that only differ in the arrangement of events. Prospect Theory (Kahneman & Tversky, 1979) makes differing predictions. Here, later events within experiences are evaluated using earlier events, i.e. so called reference points. Namely, subsequent events are perceived as gains or losses compared to the reference point, with losses of equal absolute value being higher weighted than gains. Applied to our two lunch experiences, Prospect Theory predicts a preference for the increasing sequence of the two experiences: Although the first meal in the increasing sequence (bread) has a lower value than the initial steak in the decreasing sequence, consuming steak on day two corresponds to a gain compared to stale bread, while in the decreasing sequence consuming bread on day two is a loss with the same absolute value as the gain. Hence, although both experiences are identical, depending on ordering, decision makers once perceive the second as loss and once as gain. According to Prospect Theory, losses are higher weighted than gains; thus, a

[7] This chapter deals with experiences and sequences which get better or worse. Within this study, *sequence* describes a specific type of experiences. To standardize the wording, any improvement of an experience or a concrete sequence is labeled as *increasing* while all deteriorations are labeled as *decreasing*.

decision maker will choose the increasing sequence. To sum up, when evaluating experiences using Prospect Theory in contrast to traditional rationality, the timing of events. i.e., the temporal order of events within the experience is important.

Both theories describe human behavior independent of the specific application. That is, they can be applied to any type of prospect, be it experiences, abstract alternatives or outcomes of games. Aside from these theories, behavioral anomalies have been observed, which are characteristic for experiences consisting of different events. Here, typically the complexity of the decision task based on attention, storage mechanisms and memory is considered. The basic idea is that processing of new experiences results in a mental representation of these experiences. This representation is not always adequate due to cognitive limitations or attention phenomena. Events vary for example concerning their importance for the individual. However, this imperfect representation forms the basis for subsequent valuation processes. Thus, complex valuation decisions over experiences are prone not to reflect real preferences. Relevant behavioral anomalies in the context of experiences are the peak-end-rule (Kahneman et al., 1993) and primacy/recency effects (Baddeley et al., 2009).

According to the peak-end-rule (Varey & Kahneman, 1992), the peak and the final event of an experience influence the evaluation more than all other events in the experience. While decision makers hardly judge how long an event lasted, they have a good judgment concerning the intensity of the event. In consequence, decision makers evaluate experiences by averaging the last event and the event with the highest intensity. Applied to the increasing and decreasing experience of meals, the peak-end-rule predicts a preference for the increasing experience ending with the steak on day two.

The peak-end-rule has been observed in several studies. Kahneman et al. (Kahneman et al., 1993) investigated preferences over two painful experiences. In both experiences, subjects had to put a hand in a basin of water of 14° Celsius for one minute. One of the experiences finished with 30 seconds of additional pain in slightly warmer water. The majority of subjects preferred the longer experience with more overall pain over the short experience with less pain, as predicted by the peak-end-rule. The peak-end-rule also occurs during displeasing tasks. In an experiment

(Finn, 2010) subjects had to learn two lists on difficult English-Spanish translations. To one of the equally long lists, easy vocabulary was added. Following the peak-end-rule, the majority of students favored learning the longer list with the more pleasant end.

The peak-end-rule predicts preferences over experiences with different ends. It does not differentiate if experiences differ except for their peaks and ends. Serial position effects, namely primacy and recency effects, consider this (Deese & Kaufman, 1957; Glenberg et al., 1980). According to the primacy (recency) effect, earlier (later) events have the highest impact on the evaluation of the experience (see Hastie & Park, 1986 for an overview). Serial position effects are the result of memory limitations: first events are stored in short-term memory earlier, compared with events presented afterwards. As the capacity of short-term memory is very limited, the representation of events becomes fragile after a short time delay. However, the events experienced earlier can enter long-term memory. I. e. earlier events are compared with later events during their presentation. This retrieval for comparison strengthens the representation and moves earlier events into long-term memory (Atkinson & Shiffrin, 1968). The earlier the event is experienced, the better is its representation in long-term memory as their storage process is not interrupted by preceding items. After the perception is finished, earlier events (primacy effect) are better represented due to their entering of long-term memory, whereas final (recency effect) events are better accessible because they are still present in short-term memory. Thus, first and final events stronger influence retrospective evaluations of the experience as a whole. From these theoretic considerations, it is not clear which effect, primacy or recency, stronger influences retrospective evaluations of experiences. Hence, no predictions concerning the increasing and decreasing experience of meals can be made.

Related to primacy and recency effects, the importance of memory limitations when deciding over experiences has been confirmed in an experiment on several consecutive tastings of wine (Mantonakis et al., 2009). Advanced oenophils and beginners tasted five wines subsequently. Beginners behaved in line with the primacy effect whereas oenophils followed the recency effect. Beginners compare the first wine with the second, their resulting favorite with the following wine and so on. Oenophiles search more persistent for the best wine and dispose of a higher capacity for the searching process until all options have been presented. Thus, last options get higher attention from the oenophils. Aside on background knowledge, the occurrence of primacy or recency effects depends on the valence of the objects (Demaree et al., 2004; Li & Epley, 2009). If subjects are confronted with a sequence of negative items, the negativity of first items gets reduced whereas the negativity of last items is still very present yielding to a primacy effect. This works the other way round for positive items. Positivity is present for last but reduced for first items, so that subjects remember and chose later items. To sum up, know-how as well as negative valence of events/items foster primacy whereas positive valence supports a recency effect.

It is not clear how well theoretic models, rationality assumptions or Prospect Theory, and anomalies, peak-end or primacy/recency, predict human behavior. The question is especially interesting, as – to the best of our knowledge – no study on the anomalies exists, which includes real consequences, i.e. stated preferences do not yield any subsequent monetary rewards. The studies of Kahneman et al. (Kahneman et al., 1993) and Finn (Finn, 2010) are no exception. They announce consequences, but do not realize them. Instead, subjects indicate their preference (Finn, 2010; Kahneman et al., 1993; Mantonakis et al., 2009), rate the options (Diener et al., 2001; Do et al., 2008) or compare them to other experiences (Redelmeier et al., 2003).

Analyzing the impact of real consequences is important. Recent neuro-economic studies show that hypothetical decisions have higher processing complexity and need more cognitive resources than real decisions (Morgenstern et al., 2013). Hence, it is unclear whether primacy/recency effects that are the result of cognitive limitations, are the result of hypothetical and therefore artificial questioning and whether they still occur in scenarios with real consequences. We expect that introducing real consequences reduces anomalies related to extensive processing.

To answer these questions, we conducted an experiment using the cold pressor test (Hines & Brown, 1936). Subjects experienced a sequence of increasing temperatures (from 4° C via 8 ° C to 12° Celsius) and a decreasing sequence with the same temperatures in inverse order. Afterwards, we asked subjects to state their preference using a rating scale having no consequence and we asked for their willingness to pay to go through the favored instead of the other experience again. This procedure includes real losses of money. The individual choice was realized in the end of the experiment. For our experiment, predictions are slightly different than for the increasing and decreasing lunch experience as in our experiments subjects experienced both sequences after another. According to predictions for a rational decision maker, subjects should be indifferent as both sequences consist of the same temperatures; Prospect Theory makes different predictions depending on the experience which is perceived as reference point, as we will show. The peak-end-rule clearly favors the increasing sequence. All temperatures and the peaks in both sequences are identical, but the increasing sequence with decreasing pain has the better end. The primacy effect finally favors the sequence experienced first: Both sequences have a negative valuation, resulting in subjects who relativize the earlier sequence while the elements of the later sequence are still present in short term memory. We do not expect a recency effect to occur as all events and the overall experience both have negative valence. In addition, neither theory predicts any difference between the questioning with and without monetary consequences.

We find that consequences strongly influence decisions. As soon as real consequences are added, anomalies are significantly reduced: subjects showed a clear preference for the increasing sequence using a rating scale which confirms peak-end-rule and primacy effect. However, several of them are indifferent when they face real consequences, a rational reaction. We analyzed the data considering the order in which the increasing and decreasing experiences are presented before subjects have to state their preference. We find that even without real consequences, neither of the anomalies explains behavior. Subjects who first face the increasing experience always favor it. Half of the subjects who start with the decreasing experience also prefer the increasing one and the other half prefers the decreasing experience. This results in a behavioral pattern not compatible with either the primacy effect or the peak-end-rule.

As an attempt to explain our observations, we use Cumulative Prospect Theory (Tversky & Kahneman, 1992) to derive a reference point model. The central idea of our model is that the first temperature experienced forms the individual reference point for the subsequent pain levels. In line with Cumulative Prospect Theory, these pain levels are interpreted either as gains or losses. Following this model, all results we observed are predicted – if we assume that the first event of both increasing and decreasing experiences marks the reference point for the subsequent events.

The remainder of this paper is structured as follows: in Section 2 we describe the experiment in detail, first the design, afterwards the procedure. Section 3 contains the results and Section 4 is the presentation of our reference point model, divided in an explanation for each of the two possible preferences. The last section concludes.

2.2 Experiment

In this section, we first describe the design of our study, before we discuss our experimental procedure.

2.2.1 Design

Pain Experiences: To evaluate individual preferences relative toward rational behavior, we used the cold pressor test (Hines & Brown, 1936) to induce two painful experiences, i. e. two sequences. Each sequence includes three temperatures. The increasing sequence (1) consists of rising temperatures (4°, 8°, 12° Celsius) while the decreasing sequence (2) consists of falling temperatures (12°, 8°, 4° Celsius).

Real Preferences: Subjects had to state their general preference for one of the sequences. Afterwards, subjects indicated their willingness to pay to repeat the favored sequence instead of the unfavored one.

Hypothetical Preferences: Subsequently, subjects rated pain intensities felt during both sequences using a numerical rating scale. Integer values represented experienced pain, 0 standing for no pain at all and 10 for the strongest pain possible.

The pain exposure with the three temperatures was equally long lasting in both sequences. Moreover, the first sequence experienced was randomized in the sample. I.e., a random draw decided whether subjects first experienced the increasing or the decreasing sequence. A rational decision maker has no preference for either sequence and overall preferences within the group should not correlate with the order in which the sequences were experienced. However, what can be expected according to the peak-end-rule (1) and the primacy effect (2) is different. The utility of the two sequences varies, according to the peak-end-rule (1): Whereas the peak is identical in both sequences (12° C as positive experience, 4° C as negative experience), the end is better for the increasing sequence (12° instead of 4° C). Predicting preferences in line with the primacy effect, we expect subjects to have a higher utility from the first sequence, no matter which one that is: As all experiences have a negative valence, subjects should favor earlier negative experiences over later ones: These early experiences were already removed from short-term memory and get relativized during the later experiences, while the pain experienced in the latter sequence is still very present. Hence, individuals following the peak-end-rule prefer the increasing sequence while those influenced by a primacy effect simply favor their first experience. Neither choice is conform to a rational decision maker who is just indifferent.

2.2.2 Procedure

The experiment began with a pain experience and included real and hypothetical preference elicitation (figure 7). 83 subjects participated in the experiment taking 30 minutes which was realized at the University in Magdeburg. One week before the experiment was conducted subjects were paid 12 Euros for participation (Thaler & Johnson, 1990). Every step within the experiment was induced by means of written instructions, be it the pain experience or the preference requests.

Pain Experiences: The order of painful sequences, either starting with the decreasing or the increasing one, was associated randomly. Within the sample, there was also a randomization of hands to immerse first, either the left hand or the right hand. To provide an identical start situation and to avoid effects of habituation, subjects switched hands and paused for 30 seconds before continuing with the second sequence. Subjects neither knew immersion durations nor temperatures of the water

Real Preferences: First, we asked participants to state their preference for one of the sequences. Four of them were indifferent and thus continued immediately with the pain intensity ratings for both sequences. All others subsequently gave their real preferences using willingness to pay. Our approach was derived by the method presented by Holt and Laury (2002). Each subject was confronted with 25 choices: the first alternative was the unfavored pain sequence whereas the second alternative was the favored one coupled with a certain amount of money. This amount rose from 0.00 until 5.00 Euros in steps of 0.20[8]. Thus, each subject could indicate how much money she found the favored sequence to be worth paying for. This procedure was enlarged for participants who had stated a preference but indicated no willingness to pay at all. To make sure that these subjects are not cheating, we coupled the amounts of money with the other sequence, which originally was indicated to be unfavored. In this control procedure, no subject indicated a positive willingness to pay which confirms that every subject had indicated his real preference.

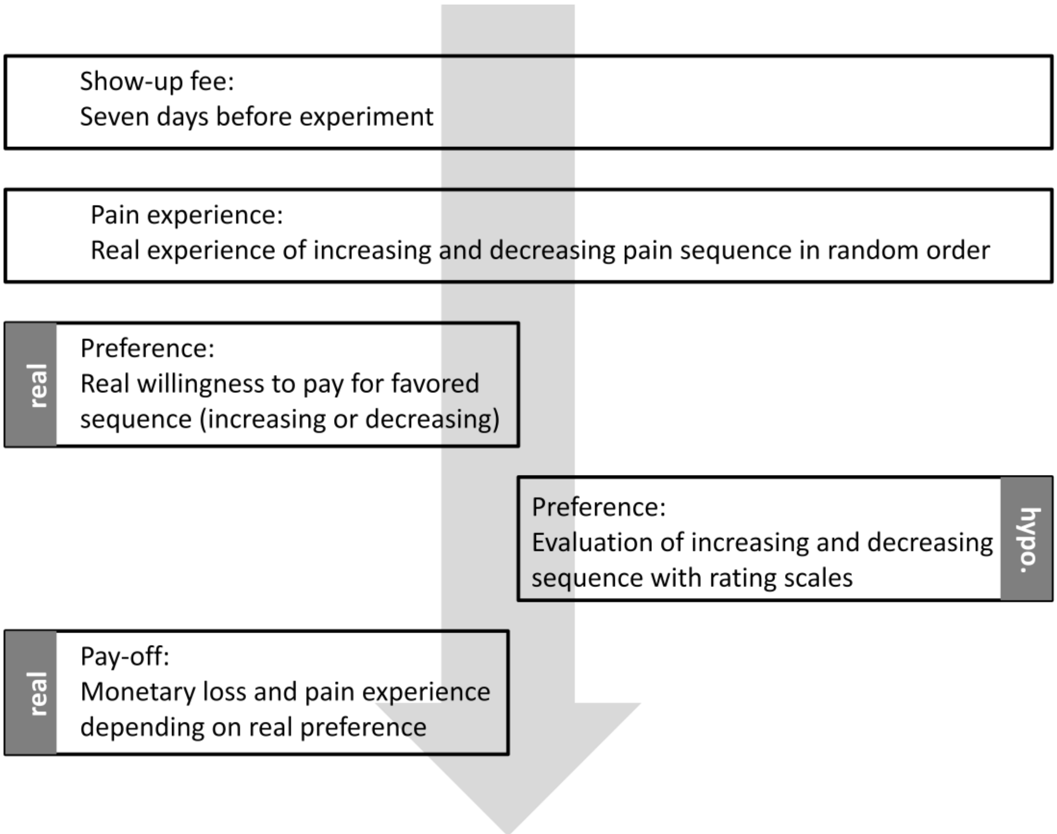


Figure 7 Experimental Procedure

[8] Appendix B presents the 25 choices in detail

Hypothetical Preferences: The experiment continued with two numerical rating scales asking to assess the experienced pain intensity during the two sequences. We ensured every subject to be aware, that these ratings neither have consequences in terms of the monetary payoff nor concerning the choice of the sequence for the re-experience.

Pay-off: After the elicitation of real and hypothetical preferences, we continued with the repetition of one of the painful sequences. Subjects who were indifferent between them in the beginning and thus did not indicate their willingness to pay repeated a sequence determined randomly. For all others we randomly chose one of the 25 choices and realized it. This means that subjects who had selected the unfavored sequence repeated this, while subjects who chose the favored sequence repeated that and additionally paid the corresponding amount of money to us.

2.3 Results

We investigated preferences over two painful sequences elicited with and without real consequences. A rational decision maker should indicate preferences neither in line with the peak-end-rule favoring the increasing sequence nor with the primacy effect contributing to preferences for the first experience.

To make the results comparable, answers on both questionnaires, willingness to pay and the rating scale, are presented normalized. Therefore, we divided all values through their theoretical maximum that is five for willingness to pay and ten for the rating scale. A positive sign indicates an effect-conform behavior: for the interpretation of data in the context of the peak-end-rule, it represents preferences for the increasing sequence, for the primacy effect it represents the favoring of first sequences.

2.3.1 Peak-End-Rule

We first analyze whether the observed behavior is in line with the peak-end-rule, i.e. whether subjects favor the increasing sequence. In both questioning procedures, hypothetical and real, the averages over all preference values lie around 0.10 (hypothetical: 0.10; real: 0.11); standard deviation is 0.25 (hypothetical: 0.24; real: 0.25) (compare figure 8) supporting the peak-end-rule. Hence, at first sight, a small tendency towards the peak-end-rule exists. This observation can be confirmed with a binomial test: significantly more subjects indicate a preference for the increasing sequence using the rating scale (p = 0.004).

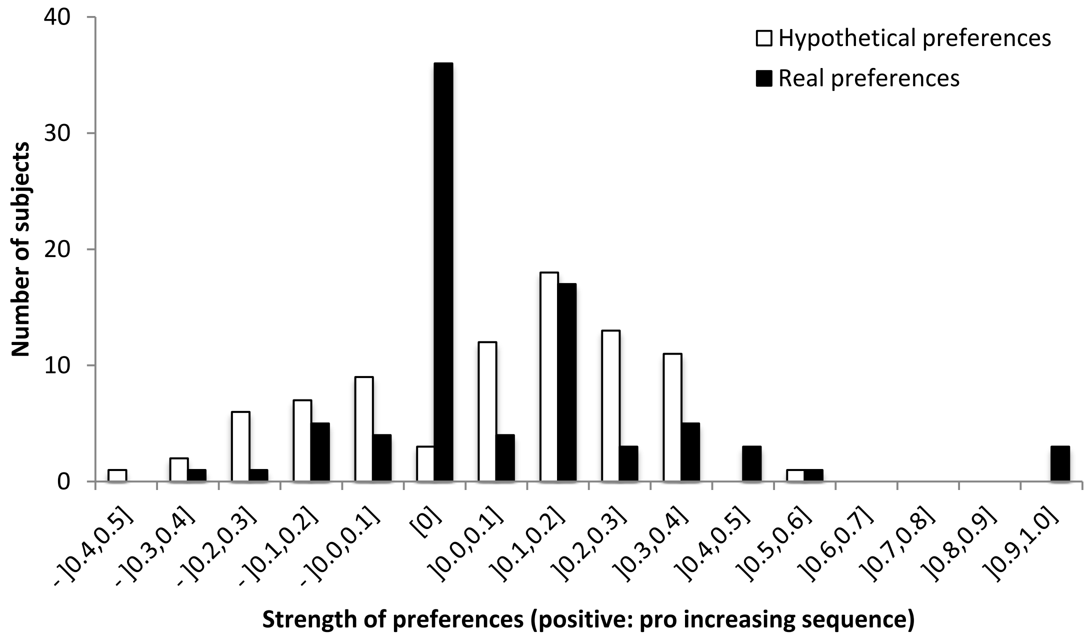


Figure 8 Real and hypothetical preferences related to the peak-end-rule

Number of subjects per stated strength of hypothetical and real preference (numerical rating scale and the willingness to pay). To easily compare the results from the rating scale with those from willingness to pay, the figure presents normalized values. Therefore, every value from the rating scale and willingness to pay was divided through their maximum (10 and 5 Euros) and included in the figure. (We calculate hypothetical preferences as difference from the specified pain intensity for the last sequence minus pain intensity for the first sequence in the corresponding questionnaire)

However, the answers to real and hypothetical questioning procedures differ in the number of indifferent subjects. 3 out of 83 subjects (4%) indicate indifference in the hypothetical questionnaire, i.e. ratings are equal for both sequences, whereas 36 subjects (43%) are indifferent on the willingness to pay questionnaire. As a result of this difference, medians of hypothetical and real preferences differ: the median for hypothetical preferences is 0.20 while it is 0.00 for real preferences. We aim to confirm this result statistically with a sign test but face the following problem. The 36 observations of indifference without any willingness to pay are ties. This means that depending on the decision to allocate them either to the values below or to those above the median, our results vary. However, we think that the high number of subjects unwilling to pay (1), the median of 0.00 (2) and the facts that subjects have preferences for the increasing and also for the decreasing sequence (3) are convincing: our participants behave rational for preference elicitations followed by real consequences by not favoring any of the two sequences. In the hypothetical procedure with the rating scales, subjects behave in line with the peak-end-rule and hardly demonstrate indifference between the increasing and the decreasing sequence.

Hence, at first sight the peak-end-rule seems to capture the observed behavior quite well. However, a look at the number of subjects being indifferent between both sequences shows that subjects answering the willingness to pay questionnaire are indifferent between both alternatives and only the hypothetical rating scale questionnaire yields a preference for the peak-end-rule.

2.3.2 Primacy Effect

Next, we investigate the observed behavior with respect to the primacy effect. That is, we analyze whether subjects show a preference for the sequence experienced first. We find for both questionnaires that subjects slightly favor the first sequence: the average value for the willingness to pay questionnaire is 0.09 and 0.13 for the rating scales, in favor of a primacy effect (standard deviations 0.25 and 0.23, respectively) (compare figure 9).

2. Do people have a preference for increasing or decreasing pain?

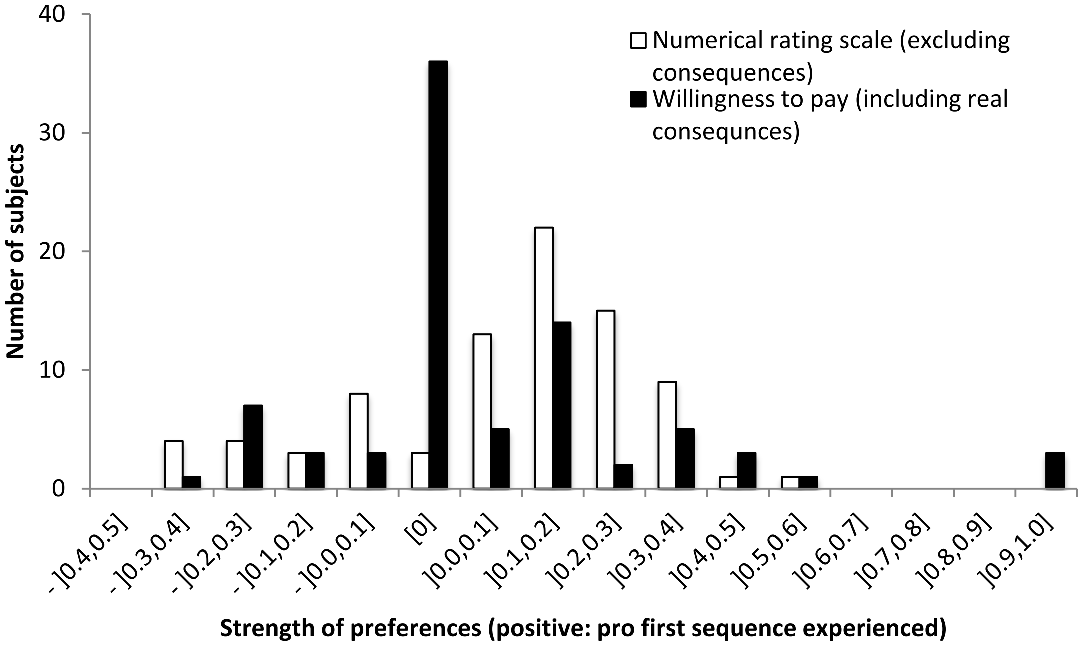


Figure 9 Real and hypothetical preferences related to the primacy effect

Number of subjects per stated strength of hypothetical and real preference (numerical rating scale and willingness to pay). To easily compare the results from the rating scale with those from willingness to pay, the figure presents normalized values. Therefore, every value from the rating scale and willingness to pay was divided through their maximum (10 and 5 Euros) and included in the figure. (We calculate hypothetical preferences as difference from the specified pain intensity for the decreasing sequence minus pain intensity for the increasing sequence in the corresponding questionnaire)

We now look at the hypothetical rating scale questionnaire in isolation. For the analysis of the primacy effect, we code whether subjects favor the first or the last sequence. A majority of 61 participants favors the first over the last, whereas 19 subjects prefer the last over the first sequence (3 ratings are equal for both sequences). Hence, there is a significant primacy effect for preference statements without monetary consequences (Binomial-test, $p = 0.00$).

Next, we investigate the data of the real questionnaire. 36 subjects are not willing to pay for any sequence. 33 (14) subjects pay a positive amount for a repetition of the first (last) sequence. We find that subjects neither favor the first sequence (Binomial-test, one-sided, $p = 0.039$) nor the last sequence (Binomial-test, $p = 0.000$).

Similar to our analysis of the peak-end-rule, we find a primacy effect in the answers to the hypothetical rating scale questionnaire. This effect vanishes as soon as monetary consequences are involved.

2.3.3 Combined view on Peak-End-Rule and Primacy Effect

Our experiment shows that the majority of subjects has a willingness to pay of 0 for experiencing a specific sequence. However, in the beginning of the willingness to pay procedure, we asked all subjects which sequence they favored. Here, only four were indifferent, all others had a preference for one of the sequences. In table 2 we compare these preferences to the answers in the rating scales of the subjects. The majority of subjects had – even if their willingness to pay (WTP) was zero – a preference in line with the values specified using the rating scale. Only one (two) subject(s) specified no preference in the willingness to pay procedure, while preferring the increasing (decreasing) sequence according to the rating scale. Another six subjects stated other preferences in the beginning of the willingness to pay procedure than they had according to their answers in the rating scales. All other 74 subjects had preferences consistent with the ratings they specified.

			sequence first seen		**sum**
			increasing	**decreasing**	
sequence preferred	**rating favors increasing**	WTP>0, preference consistent	22	10	32
		WTP=0, preference consistent	14	8	22
		WTP=0, no preference	1	0	1
	rating favors decreasing	WTP>0, preference consistent	0	10	10
		WTP=0, preference consistent	1	8	9
		WTP=0, no preference	0	2	2
	indifference	WTP>0, preference consistent	0	1	1
	others		0	6[9]	6
sum			***38***	***45***	***83***

(subjects with willingness to pay = 0 in brackets)

Table 2 Analysis of willingness to pay related to primacy effect versus peak-end-rule

[9] 3x Rating pro decreasing, WTP and preference pro increasing; 1x Rating pro decreasing, WTP pro decreasing but no preference; 2x Rating shows indifference, preference pro decreasing

The interesting part of table 2 is that it contrasts the information which sequence was presented first and which sequence the subjects favored. With only one exception, subjects who experienced the increasing sequence first also prefer it. In contrast, half of the subjects who started with the decreasing sequence prefer the increasing sequence, while the other half opts for the decreasing sequence.

2.4 Discussion

The analysis of the data shows two things: First, as soon as subjects face costs for their preferences, an increased number of them indicates a willingness to pay of 0. Second, if stating the preference is costless, two anomalies are necessary to describe observed behavior. Namely, if peak-end-rule and primacy effect make the same predictions, all subjects opt for the predicted alternative. This is the case if subjects experience the increasing sequence first. This is different as soon as subjects experience the decreasing sequence first: here, peak-end-rule and primacy effect favor opposing outcomes, so that one part of the sample behaves according to the primacy effect, the other part according to the peak-end-rule. We believe that our observation that half of all subjects follow the peak-end-rule and half behave according to the primacy effect is not a fix rule. The distribution may depend on the corresponding decision situation. Hence, the effect can give an indication what outcomes are possible but not with which probabilities the outcomes will occur.

Several of the studies we discussed in the introduction ensure the counterbalancing of treatments (Kahneman et al., 1993; Li & Epley, 2009; Finn, 2010; Diener et al., 2001): experimenters varied the order in which subjects made their experiences. However, this does not control for the biases we observed. Without separating our data based on the counterbalancing, we would still see both effects. Hence, we believe that experimental results always have to be analyzed according to peak-end-rule and primacy effect. However, to the best of our knowledge, no other study ever did this. This is especially important, as “the carriers of value or utility are changes rather than final asset positions” (Kahneman & Tversky, 1979, p. 273). Cumulative Prospect Theory can capture this. In particular, we show that simply applying Prospect Theory with different reference points suffices to explain the results without having to consider any (additional) behavioral anomaly like peak-end-rule, primacy or recency effect.

2.4.1 Prospect Theory for different temperatures

Tversky and Kahnemann (1992) introduced the value function to derive utilities of different prospects:

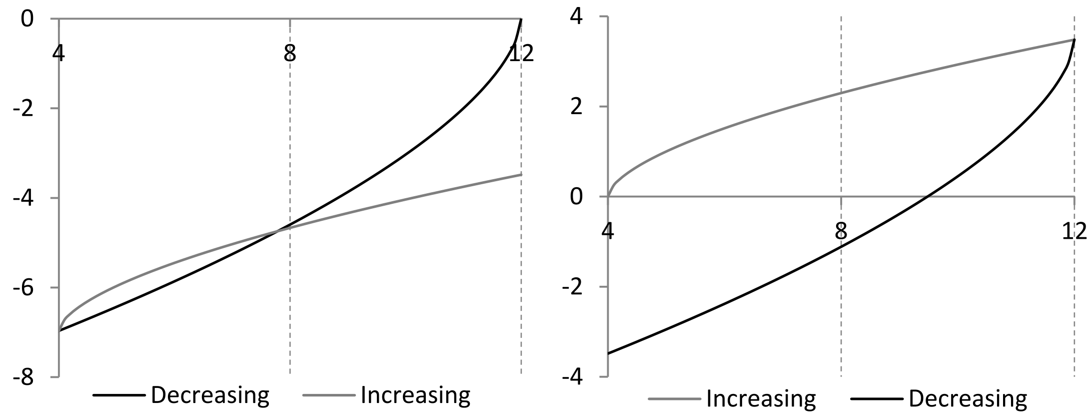
$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\beta} & \text{if } x < 0 \end{cases}$$

Their value function $v(x)$ assumes that the reference point lies at 0. However, in our experiment with temperatures the reference point $\hat{t}$ might be any temperature. So we explicitly model the reference point leading to the value function $v'(t)$:

$$v'(t) = \begin{cases} (t - \hat{t})^{\alpha} & \text{if } t \geq \hat{t} \\ -\lambda(\hat{t} - t)^{\beta} & \text{if } t < \hat{t} \end{cases}$$

In addition, we assume that related to our experiment a sequence of experiences consists of different temperatures the subjects face. These temperatures are measured in degree Celsius, there is one minimal temperature $\underline{t}$, a maximal temperature $\bar{t}$ and all other temperatures are equally distributed in the range $[\underline{t}, \ldots, \bar{t}]$. Namely, all experienced temperatures have a distance of $\frac{\bar{t}-\underline{t}}{n-1}$ with n is the number of different temperatures used. Accordingly, we are talking about three temperatures (4°, 8° and 12°C) with two intervals between them (4°-8° and 8°-12°C) and a distance of temperatures of 4°Celsius.

In the remainder of this section, we assume that participants experience both sequences after each other. To evaluate their experience they value it using the first temperature per sequence as reference point for all subsequent temperatures within the experience. We show that for subjects experiencing the increasing and afterwards the decreasing sequence, a clear preference towards the increasing sequence exists. This is different for subjects starting with a decreasing sequence who might prefer the one or the other sequence based on their individual value function.



(a) Experience of increasing after decreasing sequence

(b) Experience of decreasing after increasing sequence

Figure 10 Model for the development of values

The x-axis contains the three temperatures of every sequence 4°, 8° and 12° Celsius while the y-axis formally describes increase (positive pitch) and decrease (negative pitch) of the negative valence the sequences contain.

Figure 10 captures the basic idea of our formal considerations and explains graphically the course of valence the two sequences contain. In figure 10a the subject experiences the decreasing sequence (black line) before the increasing sequence (grey line). Hence, the first temperature (12° C) is perceived as reference point for the rest of the first sequence. Accordingly, the two following temperatures (8° C and 4° C) are perceived as losses relative to the first temperature. After switching hands, the increasing sequence starts. For this new sequence, the first temperature (4° C) is used as the new reference point. However, this temperature of the reference point is not perceived as an experience of neutral value. On the contrary, it holds the value of the last temperature of the preceding sequence (as both have the same temperature). Both later temperatures (8° C and 12° C) are improvements compared to this first temperature. However, as temperature improvements (gains) value less than temperature declines (losses), the last temperature (12° C) of the increasing sequence always has a lower value than the first temperature (12° C) of the decreasing sequence. As we will show in the next subsection, whether the overall value of the increasing sequence exceeds the overall value of the decreasing sequence depends on the intersection of both utility functions. For the experience of the increasing sequence before the decreasing sequence (compare figure 10b) similar arguments hold. However, as we will show the value functions of the increasing and the decreasing experience never intersect. Hence, the utility of the temperatures in the increasing sequence always lies above the utilities of the corresponding temperatures in the decreasing sequence. Hence, subjects facing such experiences will always favor the increasing over the decreasing sequence.

2.4.1a Decreasing – Increasing

We first investigate what subjects choose if they experience the decreasing sequence before the increasing sequence. In the decreasing sequence, each experience is perceived as loss compared to the initial (maximal) experience $\bar{t}$, the first reference point. The value $v^{'}(t)$ of a single experience is $-\lambda(\bar{t}-t)^{\beta}$ and the overall experience of the decreasing sequence with two changes of temperatures n-1 is:

$$v'_{dec} = \sum_{i=0}^{n-1} \left(-\lambda(\frac{\bar{t}-\underline{t}}{n-1} i)^{\beta} \right)$$

(Sequence of experiences: **decreasing**-increasing)

In the end of the decreasing sequence, subjects face the coldest temperature 4°C. This temperature has the maximal distance to the initial reference point (12°C) and a value of $v'(n-1) = -\lambda\left(\bar{t}-\underline{t}\right)^{\beta}$ starting with 4°C. In the increasing sequence now all temperatures are perceived relative to the new reference point which is not 12°C anymore but 4°C. Each experience has the value $v'(t) = -\lambda\left(\bar{t}-\underline{t}\right)^{\beta} + \left(t-\underline{t}\right)^{\alpha}$ where the first summand is the value at the reference point and the second summand is the utility for the deviation from this reference point. The overall value of the increasing sequence is:

$$v'_{inc} = \sum_{i=0}^{n-1} \left(\left(-\lambda \left(\bar{x} - \underline{x} \right)^{\beta} \right) + \left(\frac{\bar{x} - \underline{x}}{n-1} i \right)^{\alpha} \right)$$

(Sequence of experiences: decreasing-**increasing**)

A subject will choose the decreasing sequence if $v'_{dec} > v'_{inc}$. This is fulfilled, if

$$\lambda > \frac{\sum_{i=0}^{n-1} \left(\frac{\bar{t} - \underline{t}}{n-1} \cdot i \right)^{\alpha}}{\sum_{i=0}^{n-1} \left(\left(\bar{t} - \underline{t} \right)^{\beta} - \left(\frac{\bar{t} - \underline{t}}{n-1} \cdot i \right)^{\beta} \right)} \geq 0$$

holds.

In our experiments, $\bar{t} = 12°C$; $\underline{t} = 4°C$ and $n = 3$ holds.

$$\lambda > \frac{4^{\alpha} + 8^{\alpha}}{2 \cdot 8^{\beta} - 4^{\beta}} = \underline{\lambda}$$

From this equation, we derive a lower bound for the parameter for the sensitivity of losses $\underline{\lambda}$. Subjects having a higher lambda will choose the decreasing sequence.

2.4.1b Increasing – Decreasing

Applying the same logic as before (see last subsection) to the decreasing sequence experienced after the increasing sequence yields different results. The value of the sequence experienced first, the increasing sequence with n-1 again reflecting the two changes of temperatures within the sequence is:

$$v'_{inc} = \sum_{i=0}^{n-1} \left(\frac{\bar{t} - \underline{t}}{n-1} i \right)^{\alpha}$$

(Sequence of experiences: **increasing**-decreasing)

As the subjects have no experience concerning cold water, it excludes a summand representing any earlier experience. The value of the decreasing sequence experienced afterwards includes a summand for the reference point, i.e. the last temperature of the first sequence (12°C). It is equivalent to

$$v'_{dec} = \sum_{i=0}^{n-1} \left((\bar{t} - \underline{t})^{\alpha} - \lambda (\frac{\bar{t} - \underline{t}}{n-1} i)^{\beta} \right)$$

(Sequence of experiences: increasing-**decreasing**)

Subjects will choose the decreasing sequence if $v'_{dec} > v'_{inc}$, i.e. if

$$\lambda < \frac{\sum_{i=0}^{n-1} \left((\bar{t} - \underline{t})^{\alpha} - (\frac{\bar{t} - \underline{t}}{n-1} \cdot i)^{\alpha} \right)}{\sum_{i=0}^{n-1} (\frac{\bar{t} - \underline{t}}{n-1} \cdot i)^{\beta}}$$

With the parameters of our experiment this simplifies to

$$\lambda < \frac{2 \cdot 8^{\alpha} - 4^{\alpha}}{4^{\beta} + 8^{\beta}} = \bar{\lambda}$$

Hence, subjects will choose the decreasing sequence if their parameter for the sensitivity of losses λ lies below the upper bound $\bar{\lambda}$. This differs from the other order of experiences. For subjects starting with a decreasing sequence high loss aversion, i.e. high λ, leads to choosing the decreasing sequence.

2.4.2 Application to parameter sets from literature

In this subsection, we apply risk parameters, i.e. α, β and λ, observed in the literature to our theoretical predictions, before we relate these results to our experimental observations. Subjects in experiments are typically risk averse over gains, i.e. $\alpha < 1$, risk seeking over losses, i.e. $\beta > 1$, and more sensitive to losses than gains, i.e. $\lambda > 1$ (see e.g. Neilson & Stow, 2002). The exact height of the parameters depends on the specific experiment.

In table 3 we summarize parameter estimates from the literature. The table only shows work, which estimates all three parameters, i.e. we exclude publications estimating a subset of parameters only. We also do not question the estimation methods used in the publications. Two observations are central: (1) in most studies α and β are identical or close to identical. (2) The parameter for loss aversion λ strongly varies across studies.

Reference	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\lambda}$
(Tversky & Kahneman, 1992)	0.88	0.88	2.25
(Tu, 2005)	0.68	0.74	3.20
(Andersen et al., 2006)	0.81	0.80	1.07
(Abdellaoui et al., 2007)	0.72	0.73	2.54
(Abdellaoui et al., 2008)	0.86	1.06	2.61
(Rieskamp, 2008)	0.91	0.91	1.00
(Harrison & Rutström, 2009)	0.71	0.72	1.38

Table 3 Parameter estimates of risk parameters from literature

The observation (1) that $\alpha = \beta$ (also see e.g. Neilson & Stow, 2002) simplifies our analysis. With $\alpha = \beta$ one can easily show that both $\underline{\lambda}$ and $\overline{\lambda}$ are monotonically increasing in α, as

$$\frac{\delta\underline{\lambda}}{\delta\alpha} = \frac{4^{\alpha}\cdot 8^{\alpha}\cdot \ln(8) + 3\cdot 4^{\alpha}\cdot 8^{\alpha}\cdot \ln(4) + 2\cdot 4^{2\alpha}\cdot \ln(4)}{(2\cdot 8^{\alpha} + 4^{\alpha})^2} > 0 \text{ and}$$

$$\frac{\delta\overline{\lambda}}{\delta\alpha} = \frac{3\cdot 4^{\alpha}\cdot 8^{\alpha}\cdot \ln(8) + 8^{\alpha}\cdot 4^{\alpha}\cdot \ln(4) - 2\cdot 4^{2\alpha}\cdot \ln(4)}{(8^{\alpha} + 4^{\alpha})^2} > 0 \text{ hold.}$$

With this and given the assumption that subjects are risk averse for gains and risk seeking for losses, i.e. $0 \leq \alpha \leq 1$, we can calculate the value range for $\underline{\lambda}$ and $\overline{\lambda}$ by inserting the minimum (0) and maximum value (1) of α into our equations for $\underline{\lambda}$ and $\overline{\lambda}$ in the preceding subsections.

Hence, subjects facing the decreasing sequence first, choose it if

$$\lambda > \underline{\lambda} \in [1, 2]$$

Subjects starting with the increasing sequence choose the decreasing sequence, if

$$\lambda < \bar{\lambda} \in [\frac{1}{2}; 1]$$

Given that subjects are more sensitive over losses, i.e. $\lambda > 1$, an assumption commonly accepted, subjects starting with the increasing sequence will never choose the decreasing sequence, as for them $\lambda < 1$has to hold. This supports the results we observed in our experiments.

For subjects starting with the decreasing sequence, making predictions is more difficult. As for $\lambda > \underline{\lambda} \in [1, 2]$subjects choose the decreasing sequence and all values are feasible because loss aversion only restricts λ to be 1 or higher, several subjects might choose the decreasing over the increasing sequence. However, what fraction chooses the one or the other sequence is difficult to predict. We can only discuss an indication: As parameter estimation studies (compare table 3) find average λ around 2, about half the subjects might have a λ below 2 and half above 2. In consequence, our experimental observation of 50% of our subjects choosing the increasing (decreasing) sequence is at least plausible.

2.5 Conclusion

We graphically and formally showed that subjects will choose different sequences if they experience both sequences in different order. The formal predictions resort to nothing but standard Prospect Theory and assumptions on the parameters widely accepted and frequently observed in the field. The behavioral results we found in our experiments are in line with these formal predictions. This compliance of behavioral results and formal predictions is especially surprising: Cumulative Prospect Theory clearly outperforms psychological anomalies as two competing anomalies, namely primacy, and peak-end, are necessary to predict behavior.

Chapter 3[10]

3. Determining risk preferences for pain

3.0 Motivation

Contrary to the assumptions of prevalent behavioral anomalies like the peak-end-rule or primacy effect, subjects in the previously described experiment demonstrate rational behavior in situations with real consequences. Transferring these results to the medical context, positive consequences for the decision situation over possible treatments can be expected. Shared decision-making of physician and patient always takes place in a real scenario including the ultimately affected patient with his specific health condition. According to the literature (Godolphin, 2009), physicians do rarely practice shared decision-making, be it because of restricted time slots or limited communicational competences. Although the patient is somehow involved in the decision process, we can hardly assume that this process is shared and includes a reciprocal communication style. Here, the patient is often passive awaiting the recommendations of his physician. The specific way how the physician presents his conclusions and recommendations can be decisive. We can hardly expect physicians to control their communication style in terms of resulting behavioral anomalies, focusing on primacy effects or related biases. However, the results of the study presented in chapter 2 show that the rational evaluation competences for example of a patient must not be underestimated but should direct the majority towards a sophisticated decision.

[10] This chapter is to a great extent based on Kroll, Trarbach & Vogt, (2011).

Behavioral anomalies are only one aspect relevant to control for if the decision situation is to be investigated. Especially in terms of resource allocation decisions, the basis is evaluating data describing the effects of optional interventions. Accordingly, the resulting amelioration of health such as a better quality of life or an enlarged life expectancy must be assessed. Therefore, different methods are available. However, besides the described willingness to pay method which is seldom utilized in this context, the approach used worldwide and with the highest acceptance is the QALY-concept. The QALY index represents information on the two components, namely quality of life and life time spent within a certain level of quality of life. The next investigation focuses on the underlying risk attitude of patients towards quality of life and life time.

The realization of this study again is done with the cold pressor test for controlled pain induction. Analogous to the experiments in chapter 1 and 2, within the setting subjects were not confronted with any wording related to medical questions.

3.1 Introduction

Quality-adjusted-life-years (QALY) are the outcome measure which is implemented in most studies on cost-utility evaluation of health care programs. The QALY takes into account the two decisive parameters relevant in terms of global resource allocation or individual treatment decisions: health-related quality of life and remaining life years. Both factors are multiplied to generate a QALY-index. A value for quality of life can be determined using standard procedures like time-trade-off or questionnaires. Thus, the value for quality of life can be implemented as a weighted factor. This is different for the second component; remaining life time is included as a linear factor. Hence, the value of a health state is linearly related to the time spent in this state. The aim of our study is to investigate the shape of the two utility functions relevant in the QALY concept, one for quality of life, the other for remaining life years.

The QALY concept bases on expected utility theory supposing a diminishing marginal utility for amelioration as well as for deterioration of health. Expected utility includes that probabilities for outcomes are available which is rarely the case in the individual situation. Another approach can reasonably be applied to describe health-related decision situations. Prospect Theory is based on uncertainty and differentiates between gains and losses. If health-related quality of life is reduced by sickness or other factors, the individual subject perceives this change as a loss. The same applies to circumstances causing a shorter life expectancy. Both, reductions in quality of life and the duration of this reduction can be categorized as losses. According to Prospect Theory the utility function is convex for losses. Hence, we can assume both utility functions for quality of life and time to be convex. While the QALY concept allows a non-linear utility function for quality of life, it assumes a linear function for time.

We apply a standard procedure from experimental economics to elicit the utility function for quality of life and time using lottery choices. Our experimental design includes real consequences from lottery decisions. For quality of life we find subjects to be risk averse which differs from the risk seeking assumption for losses made in Prospect Theory. However, the second experiment shows in line with

Prospect Theory that participants are risk seeking for varying durations of reduced quality of life, which contradicts the linearity assumption of the QALY concept.

3.1.1 Quality adjusted life years

The QALY concept is constructed related to the axiomatic structure of expected utility theory (Weinstein et al., 2009; von Neumann & Morgenstern, 1944). Pliskin et al. (1980) theoretically investigate which requirements must be postulated for the QALY to be consistent with expected utility theory. They identify three criteria that have been the conditions commonly used for several years to define the frame in which QALYs are valid. First they describe utility independence related to life years and health status. The second point is a constant proportional trade-off. It means that in a theoretical trade-off situation, one does not take into account the individual remaining life time when trading life years for some improvement in health. Finally, they assume that the individual is indifferent as to whether the result is a gamble or a sure outcome, both given in life expectancy. Here, the expected value of the gamble is equal to the sure result, implying risk neutrality for decisions about life years. However, this assumption of risk neutrality is rarely validated in any study. An exception are Miyamoto and Eraker (1985) who find risk neutrality in the mean subject but seldom in a real study participant. Later, Bleichrodt et al. (1997) demonstrate that in the medical context the postulation of risk neutrality for life years is enough to render QALYs applicable.

Several studies have investigated the QALY-assumption of risk neutrality towards life years but rarely find supporting results. McNeil et al. realized interviews on risk posture over time with both, cancer patients during their therapy (1978) and former cancer patients (1981). In both studies participants demonstrate risk aversion over life years. Several other studies working with standard gamble or time-trade-off and healthy subjects also find risk aversion (Oliver & Cookson, 2010; Verhoef et al., 1994; Stiggelbout et al., 1994; Rosen et al., 2003). Furthermore certain factors are determined that influence the individual risk attitude. In line with Prospect Theory, subjects´ risk posture over gambles on life years switches from seeking to averse when the expected value of life years increases (Verhoef et al., 1994). Stiggelbout et al. (1994) demonstrate that patients who underwent chemo therapy are more risk averse than patients who had been in surveillance protocol only. Moreover, race, gender and education influence risky decisions (Rosen et al., 2003): African Americans, men and people with lower education are less risk averse.

3.1.2 Integration of temporal information within experiences

Besides the QALY-related research on time preferences, studies that generally investigate perception and processing of temporal information demonstrate non-normative decisions (Varey & Kahneman, 1992; Ariely, 1998; Ariely & Loewenstein, 2000; Diener et al., 2001; Fredrickson & Kahneman, 1993). It seems particularly difficult to estimate peoples´ risk attitude over life years in the light of findings repeatedly demonstrating deviations from utility maximization in terms of discomfort duration. Kahneman et al. (1993) report in their study on preferences of two painful sequences of different lengths that participants favor the sequence of longer duration and more pain, and hence less utility. In line with this experimental finding including healthy subjects, studies with patients undergoing colonoscopy demonstrate the same: longer durations of a painful procedure are preferred (Redelmeier & Kahneman, 1996; Redelmeier et al., 2003). Both authors refer to the better ending of the sequences as reason for such non-normative preferences.

Apart from the relevance of specific sections of an experience like the end, Schreiber and Kahneman (2000) report additional findings concerning the temporal integration within painful experiences. Whereas a multiplicative relation of the duration and the quality of an experience is the logical combination, the authors elicit an additive extension effect. It means that experiences are mainly encoded in terms of their value and temporal information is associated additively. Kahneman (2000) adds that temporal information is not memorized in every situation or stored separately. Hence, the relevance and retrieval of temporal facets of an experience are reduced. The scope of this reduction seems to depend on the attentional focus and the study design with its measure of global evaluation (Ariely & Loewenstein, 2000; Ariely et al., 2000).

3.1.3 Utility functions over pain

Economic research avails itself of multi-attribute utility functions to describe tradeoffs such as in the QALY concept between the two dimensions quality of life and time. Because of the difficult operationalization of quality of life, our study uses a central component of it, pain. We aim to specify the curvature of two utility functions: over pain intensity (reflecting a reduction of quality of life) and duration of pain (reflecting life time spent in a certain health state). Pain is induced with the cold pressor test (CPT) and includes the immersion of a hand in water of painful cold temperature. This generates a deep tonic, thermal pain (Lorenz, 2002). It is a commonly applied procedure in pain research (Streff et al., 2010; Lovallo, 1975; Kahneman et al., 1993; Hines & Brown, 1936), especially for studies simulating chronic pain (Mitchell et al., 2004).

Finding an objective cardinal measure for pain intensity is difficult since pain perception is a complicated process influenced by various aspects like attention and distraction or gender (Keogh et al., 2000; Keogh & Herdenfeldt, 2002). However, the same is true for the measuring of the perceived economic value of any object. Therefore, economic theory uses objective measures and transfers them into utilities, which reflects individuals´ perception of value. The economic value of a good is determined using monetary values, a cardinal measure (i. e. the value difference between \$5 and \$10 is equal to the difference between \$10 and \$15). However, the difference in utilities is not equal (u(10)-u(5) > u(15)-u(10)). This is captured by a concave utility function for money. We apply the same argument for the individual perception of pain. One frequently used objective measure for pain is temperature, also implemented in our study to determine the utility function for pain. For temperatures the same characteristics apply as for different amounts of money and its divergent utilities: the difference between 5° and 10° Celsius is equal to the difference between 10° and 15° Celsius though the utilities for the temperatures in the cold pressor test are not.

Besides hand immersion into cold water there are other objective measures for pain, such as intensity of electric shocks (Berns et al., 2008) or aversive sounds (Schreiber & Kahneman, 2000). For our experiment we use temperature, but the general setup can be applied to any other measure. Thus, in any study on preference or valuation, objective measures like amounts of money, temperatures or intensity of electric shocks are presented to a decision maker. However, his decision is based on the individual utility of the object or experience to evaluate.

3.1.4 Real consequences for preferences

The empirical research on preferences in health contexts relies on questionnaires and hypothetical choice situations. However, the utility function elicited from stated preferences can vary between hypothetical and real choice situations, as shown for the utility function for money (Holt & Laury, 2002, 2005). Therefore, it seems necessary to apply experimental methods, in which subjects face real consequences of their choices to investigate health-related decision making. In our experiment we include the Holt and Laury procedure (2002) to elicit both preferences for different pain intensities and durations. Modifications are realized so that the consequences of the original lotteries are replaced with pain intensity and time. Our treatment concerning pain intensity varies the temperatures of the cold water bowl and thus the experienced level of pain. In the second treatment on pain duration, the lottery outcomes in the Holt and Laury procedure are related to different immersion durations. Using both, the procedure from Holt and Laury (2002) and the cold pressor test, subjects face decisions about real consequences in the dimensions essential to health-related decision making.

Following the empirical evidence from research in health economics, we find that risk attitude toward pain intensity has not been determined so far. Based on results related to time preferences, we expect subjects to show risk averse behavior for decisions about time (McNeil et al., 1978; McNeil et al., 1981; Verhoef et al., 1994; Oliver & Cookson, 2010; Stiggelbout et al., 1994; Rosen et al., 2003). However, both factors under consideration, pain intensity and duration of pain, can be interpreted as a loss since it can be reasonably assumed that the higher the immersion duration or the stronger the pain, the lower the level of well-being. Prospect Theory assumes risk seeking behavior for losses (Kahneman & Tversky, 1979, 1992). Thus, Prospect Theory and the empirical findings in health-related decision making provide different predictions.

3.2 Experimental procedure and task

3.2.1 Experimental procedure

The group of participants consisted of 85 students (41 female) from different fields of study recruited using ORSEE (Greiner, 2004) and assigned randomly to the two experimental treatments. 42 subjects took part in the pain intensity treatment and 43 subjects in the pain duration treatment. Students were not informed in the invitation that it was an experiment on pain and had no prior experience with cold pressor type experiments. The experiment was conducted at the laboratory of the Department for Sensor Technology at Otto-von-Guericke-University Magdeburg in sessions with one participant each. The laboratory provided the equipment to administer the cold pressor test using four circulating coolers. These machines included a water bowl for which the water temperature could be regulated by a thermostat. Additionally, a pump guaranteed that within the bowl the temperature was the same everywhere; a thermometer was installed in a distance of about five centimeters from the immersed hand. Thus, the apparatus allowed for a high level of standardization for a reliable measurement (Mitchell et al., 2004).

At the beginning of every session the subject read an instruction with general information on the experiment (compare figure 11, 1st instruction).

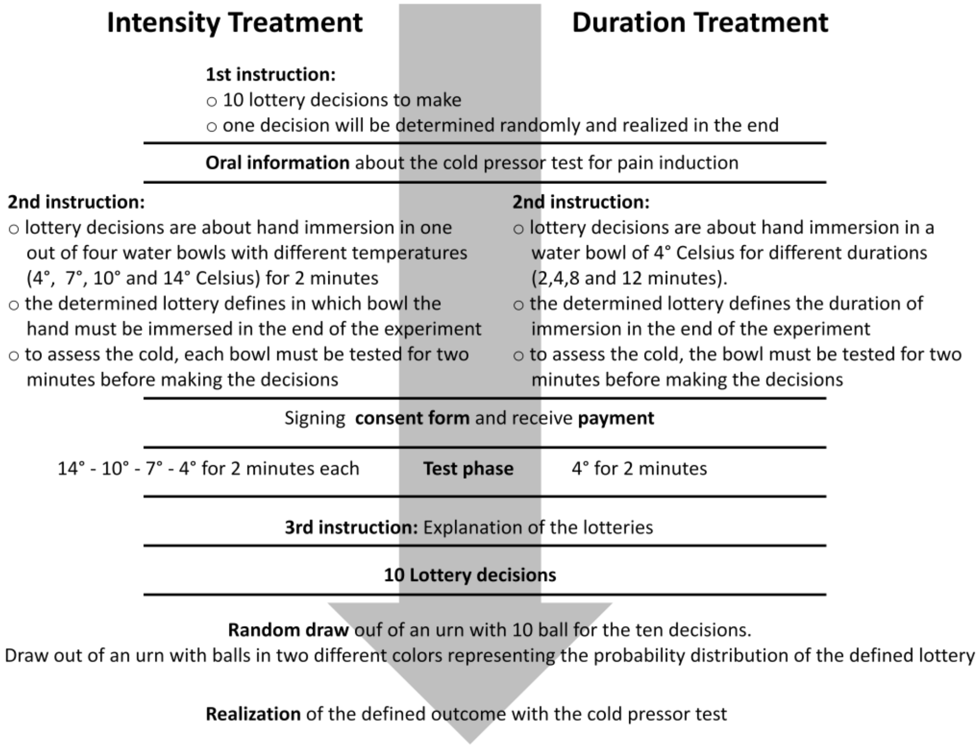


Figure 11 Comparative presentation of the experimental procedures for intensity and duration treatment

Subsequently participants were orally informed that the experiment was about pain, induced with cold water in which the hand has to be immersed. At the same time, the apparatus was presented including the display at every bowl indicating the current water temperature.

Intensity treatment: Four water bowls with temperatures of 4°, 7°, 10° and 14° Celsius were standing in line on a table. The experiment continued with the second written instruction to explain the experimental procedure (compare figure 11, 2nd instruction). If subjects agreed to participate, they signed a consent form and received 10 Euros for their participation. Note that all subjects actually participated. After the second instruction was read carefully and remaining questions were answered, the test phase began. Subjects tried each bowl for two minutes starting with the warmest and getting colder stepwise. After each bowl, there was a break of 30 seconds. The first hand immersed was determined randomly and subjects switched the hand after each bowl. The experimenter had a stopwatch to control the time and gave briefings for the participant to immerse in the next bowl, to pause or to change the hand.

Duration treatment: For the duration treatment one bowl of 4° Celsius was installed. The experiment continued with the second instruction (compare figure 11, 2nd instruction) and the consent form. All subjects signed the consent form. After the money was paid (again 10 Euros), subjects continued with the test phase. It included the immersion of one hand into the bowl for two minutes. Here again the hand was determined randomly. Except of two subjects who left at the beginning of the test phase and hence paid back the money, all subjects continued until the very end of the experiment

The next step was the 3rd instruction explaining the decision making part including the ten lotteries analog to Holt and Laury (2002). In the end of the experiment, one lottery decision was determined randomly. Therefore we took an urn with ten balls for the ten lotteries. Using another urn representing the probability distribution of the determined lottery we defined the final outcome.

Intensity treatment: Subjects had to repeat the bowl with the defined temperature. The time of immersion was two minutes again.

Duration treatment: Participants had to immerse their hand for the defined duration, 2, 4, 8 or 12 minutes. Immersion without break never exceeded two minutes. For longer durations, the time was cut into sequences of two minutes. Analog to the test phase in the intensity treatment, subjects changed the hand after every two-minutes-block and took a break of 30 seconds between blocks.

All information in terms of the exact cold pressor procedure was explained to the subjects before they made their decisions.

3.2.2 Experimental task

The ten lottery decisions were constructed analog to Holt and Laury (2002): in ten lines there was a pair of two lotteries A and B and the participant had to decide for one of them. The A and B lotteries contained the same outcome in every line (either the possible water temperatures or the immersion durations), but the probabilities were changing stepwise from 0 to 1 or rather conversely, in steps of 0.1. Lottery A presented a better and a worse option whereas lottery B contained two different moderate outcomes (compare tables 4 and 5). Most subjects switched only once between lottery A and B. In fact, they choose the moderate lottery with the intermediate outcomes at the beginning and with rising probabilities for the good outcome (warmer water temperature or shorter immersion durations), they switched to lottery A. Hence, the ten lottery decisions reflect risk seeking, risk neutral or risk averse preferences, depending on the exact lottery where subjects switch. The lottery choices a risk neutral decision maker would prefer are highlighted in tables 4 and 5, according to the expected value differences.

Although the pain intensity and the pain duration treatment are independent, in the construction of the lottery decisions we implemented a parallel construction. What generally differentiates both is that in the intensity treatment, higher values stand for a more preferable outcome (warmer water temperature) whereas in the duration treatment, higher values represent worse outcomes (longer immersion durations). Therefore, both lottery presentations are mirror-inverted concerning the point where a risk neutral decider would switch from lottery B to lottery A. Thereby, we can parallelize the allocation of lotteries in two respects:

a) First, in both treatments lottery A includes the more risky options with stronger differing outcomes whereas lottery B contains the two moderate outcomes.
b) Second, both presentations start with lower probabilities for better outcomes.

Lottery number	Lottery A		Lottery B		EV difference*
1	0.1, 14° C	0,9, 4° C	**0.1, 10° C**	**0.9, 7° C**	-2.3
2	0.2, 14° C	0,8, 4° C	**0.2, 10° C**	**0.8, 7° C**	-1.6
3	0.3, 14° C	0,7, 4° C	**0.3, 10° C**	**0.7, 7° C**	-0.9
4	0.4, 14° C	0,6, 4° C	**0.4, 10° C**	**0.6, 7° C**	-0.2
5	**0.5, 14° C**	**0,5, 4° C**	0.5, 10° C	0.5, 7° C	0.5
6	**0.6, 14° C**	**0,4, 4° C**	0.6, 10° C	0.4, 7° C	1.2
7	**0.7, 14° C**	**0,3, 4° C**	0.7, 10° C	0.3, 7° C	1.9
8	**0.8, 14° C**	**0,2, 4° C**	0.8, 10° C	0.2, 7° C	2.6
9	**0.9, 14° C**	**0,1, 4° C**	0.9, 10° C	0.1, 7° C	3.3
10	**1.0, 14° C**	**0,0, 4° C**	1.0, 10° C	0.0, 7° C	4

(EV = expected value)

Table 4 Lotteries for the pain intensity treatment

Lottery A contains the best and the worst option, i. e. the warmest and the coldest water temperatures possible (14°C and 4°C), whereas lottery B presents the intermediate temperatures (10°C and 7°C). Lottery A and lottery B are constructed with the same probabilities: low probabilities for the better outcomes (14°C and 10°C) and high probabilities for the worse outcomes (4°C and 7°C) for the first five lotteries, vice versa for the second half of the lotteries.

*Expected value difference: difference of the expected value (=expected water temperature) from lottery A and lottery B to define the lottery choice a risk neutral decision maker would choose (highlighted)

Lottery number	Lottery A		Lottery B		EV difference*
1	0.1, 2 min	0,9, 6x2min	**0.1, 2x2min**	**0.9, 4x2min**	3.4
2	0.2, 2 min	0,8, 6x2min	**0.2, 2x2min**	**0.8, 4x2min**	2.8
3	0.3, 2 min	0,7, 6x2min	**0.3, 2x2min**	**0.7, 4x2min**	2.2
4	0.4, 2 min	0,6, 6x2min	**0.4, 2x2min**	**0.6, 4x2min**	1.6
5	0.5, 2 min	0,5, 6x2min	**0.5, 2x2min**	**0.5, 4x2min**	1
6	0.6, 2 min	0,4, 6x2min	**0.6, 2x2min**	**0.4, 4x2min**	0.4
7	**0.7, 2 min**	**0,3, 6x2min**	0.7, 2x2min	0.3, 4x2min	-0.2
8	**0.8, 2 min**	**0,2, 6x2min**	0.8, 2x2min	0.2, 4x2min	-0.8
9	**0.9, 2 min**	**0,1, 6x2min**	0.9, 2x2min	0.1, 4x2min	-1.4
10	**1.0, 2 min**	**0,0, 6x2min**	1.0, 2x2min	0.0, 4x2min	-2

(EV = expected value)

Table 5 Lotteries for the pain duration treatment

Lottery A contains the best and the worst option, i. e. the shortest and the longest immersion durations possible (2 minutes and 12 minutes), whereas lottery B presents the intermediate immersion durations (4 minutes and 8 minutes). Lottery A and lottery B are constructed with the same probabilities: low probabilities for the better outcomes (2 minutes and 4 minutes) and high probabilities for the worse outcomes (12 minutes and 8 minutes) for the first five lotteries, vice versa for the second half of the lotteries.

*Expected value difference: difference of the expected value (=expected immersion duration) from lottery A and lottery B to define the lottery choice a risk neutral decision maker would choose (highlighted)

3.3 Results

The Holt-Laury procedure as used in both experimental treatments is designed to elicit individual risk preferences. Following the differences in expected values of the two lotteries, subjects are expected to switch from lottery B to lottery A, which all participants do except of two people who chose lottery A only in the intensity treatment. For both treatments, subjects' individual risk preferences can be calculated using this switching point. For the purpose of this analysis we classify subjects´ behavior only as risk averse, risk neutral and risk seeking.

In the pain intensity treatment, subjects perform decisions about pain intensities. The median of subjects switches from lottery B to lottery A between a probability of receiving the lower pain intensity of .5 and .6 (compare figure 12). Following the differences in expected values, the median observation is classified as risk averse behavior. Comparing risk averse with risk seeking subjects, we find significant risk averse behavior for lotteries on pain intensity (Binomial-Test, 5%-level).

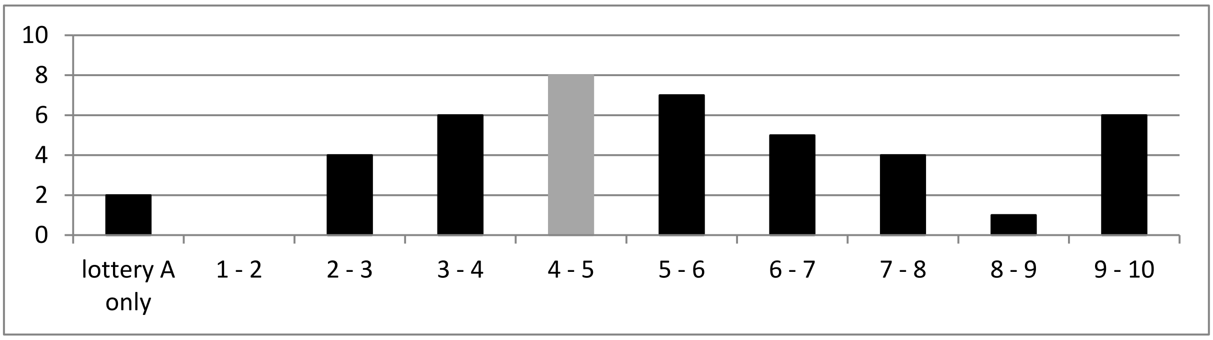


Figure 12 Switching points in intensity treatment

Number of subjects switching between the different lotteries in the intensity treatment. Switching between lotteries 4 and 5 represents risk neutrality (gray), a later switch stands for risk aversion.

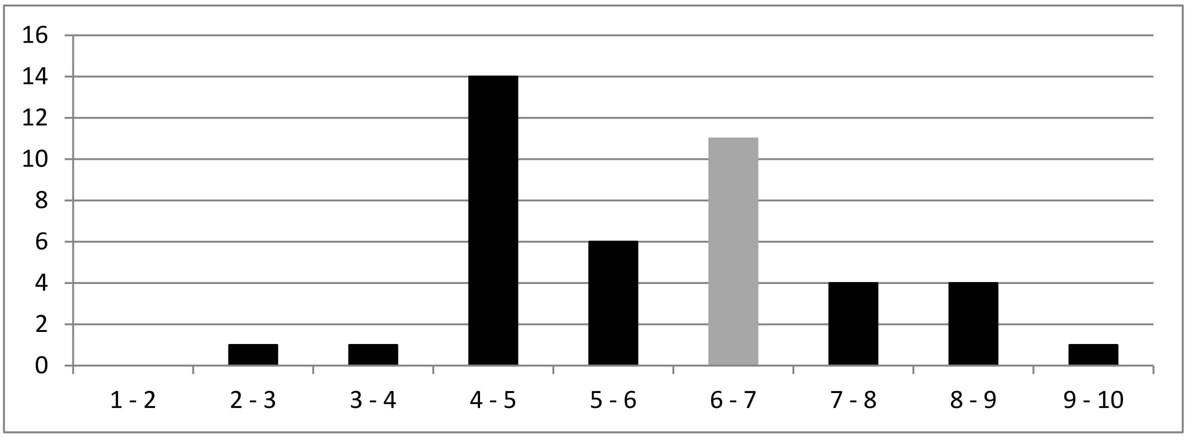


Figure 13 Switching points in duration treatment

Number of subjects switching between the different lotteries in the duration treatment. Switching between lotteries 6 and 7 represents risk neutrality (gray), a later switch stands for risk aversion.

The lottery outcomes in our experiment can be categorized as losses. Subjects face pain in any case, in the first treatment varying only in intensity. According to Prospect Theory, people behave risk seeking in decision situations that refer to losses (Kahneman & Tversky, 1979). However, behavior of the subjects in this experiment does not show risk seeking preferences for pain intensities.

In the second treatment concerning pain duration the median of the sample switches between item 5 and 6. Thus, a change from lottery B to lottery A is preferred though A has a probability of 0.4 for the longest immersion duration. According to the expected value differences, the median observation is risk seeking. We compare the risk averse group with the risk seeking one which shows that our sample is significantly risk seeking (Binomial-Test, 5%-level).

Again in the second experiment we are working with outcomes that must be perceived as losses. Subjects have to spend different durations immersing their hand in cold water which induces a tonic pain. Consonant with Prospect Theory our subjects behave in a risk prone fashion for the described decision situation on losses. On the other hand, for example the mentioned empirical study presented from Oliver and Cookson (2010) demonstrates risk averse behavior for decisions on life years. Consequently, the results of our pain duration experiment cannot easily be integrated into the results of other investigations. The question comes up how effective a real scenario with instant consequences is in comparison to hypothetical settings used for example by Kahneman and Tversky (1979) or in the other empirical studies (McNeil et al. 1978, Stiggelbout et al. 1994, Oliver & Cookson 2010). In general, it seems important to focus on this core difference as its influence might be underestimated. A significant indicator therefore is the difficulty of combining our results with the existing literature. We must increase research that includes real scenarios for example using experimental approaches.

Two central aspects of the QALY concept are key variables in our experiment: limitations in quality of life and remaining life expectancy. What we demonstrate in our study is that people are not risk neutral when it comes to limitations in quality of life; in fact they are risk averse. Additionally, when a temporal factor is included, subjects behave in a risk seeking manner. These findings clearly demonstrate that people are not risk neutral when it comes to their health. Hence, the QALY assumption of linear time preferences is hard to defend. Our scenario includes both, limitations in quality of life and different time durations. What we did not find either in the first or the second experiment is risk neutrality. On that score an adjustment of the QALY seems inevitable if we want to represent how people really interpret situations. To do so, more experiments are necessary to better understand the decision making process in health-related decision making. Experimental analyses must be central in these research questions to allow real consequences within the setting.

3.4 Conclusion

In our study we use two treatments to investigate risk preferences for decisions about pain intensity and pain duration. We realize the experiments using the cold pressor test as a standard method for investigating pain perception. This method allows us to elicit risk preferences involving pain intensity and duration using choice scenarios where subjects face real consequences from their decisions. To investigate risk preferences we use two similar decision sheets designed like those of Holt and Laury (2002): one for different temperatures, the other for different immersion durations. We find that people are risk averse for pain intensity and risk seeking for pain duration.

This result is relevant in terms of the central QALY assumption of risk neutrality for life years. It shows that subjects´ behavior is not in line with linear time preferences; hence, this simplistic assumption cannot be confirmed. Additionally, our treatment for pain duration shows risk seeking behavior which cannot be integrated into the empirical findings where subjects are identified as risk averse.

Differing from Prospect Theory and its assumption of risk seeking behavior for losses, our subjects show risk averse preferences when it comes to stronger pain intensity. Pain can be categorized as a loss and thus our results demonstrate a case where people deviate from the standard risk seeking assumption for losses.

4. Conclusion

This dissertation investigates several aspects in the context of medical decision making. The three presented experiments focus on situations relevant in the health care sector occurring on a daily basis in clinical routines. The investigated situations belong to the micro level involving the care giver as well as the patient. Both, patients and physicians, are obliged to make decisions in the course of a treatment plan. Moreover, patients and physicians have preferences over these alternatives, be it in terms of the time a treatment takes, the money to be spent for it or the associated reduction of well-being during the therapy. However, the unique situational factors as well as the individual differences of every person compound the definition of a behavioral theory describing a roadmap how individuals come to an opinion and finally make their decision.

In chapter one, the conducted experiment included a decider either with or without medical expertise who had to prioritize over five individuals in need. Three different groups of parameters formed the basis to come to an allocation decision. These groups of parameters were socio-economic attributes (1), information concerning the individual well-being or pain (2), and finally the willingness to pay of every needy to avoid the reduction of well-being as well as the information how much the prioritization of each of them costs (3). Whilst prioritization decisions based on the socio-economic information such as academic background of local origin could be categorized as more irrational allocation criteria, the criterion of pain intensity can be described as rational. The last group of parameters (3) is more heterogeneous but refers to monetary entities: obviously, the costs to prioritize the needy is essential as the overall budget the decider has for her disposal is restricted and too small to help all others in need. The individual willingness to pay to avoid the pain doses contains information about the expected level of sufferance from the pain dose allocated randomly. It can be assumed that every subject from the patient group perceives the pain from the cold pressor test as differently bad. Accordingly, there are differences in the preparedness to pay for avoidance. With willingness to pay the deciders can

rely on an additional parameter describing the need for prioritization of the affected individual, comparable over the whole group of needy subjects.
The main outcome of the first experiment is that the deciders indicate those criteria as relevant for them which they really focus on, controlled with the data from Mouse Lab: socio-economic data are irrelevant to the deciders, no matter if this decision maker is a prospective physician or a decider without any medical background. The main difference between the two groups of decision makers is the implementation of willingness to pay information to come to a prioritization decision. Prospective physicians significantly stronger include subjective information for prioritization urgency into their decision. According to our definition of efficiency, their general allocation pattern is more efficient in comparison to the non-medical decision makers.

The presented first study includes a decision scenario of high complexity for the decider. Eight parameters over five individuals in need are available for usage. Moreover, the decider has to specify her strategy to come to a prioritization decision. In clinical routines as well as with the family physician also complex decision scenarios must be handled. Thereby, the complete set of decisive parameters is seldom transparent to the patient and also the physician might not be aware of the restrictions, attitudes and prejudices she is subject to. In the conducted experiment, the whole set of potentially decisive parameters is obvious and the decision process can be followed with Mouse Lab. Only by means of such experiments, there is a real chance to enlarge the process of a deeper understanding what really guides the physician in the concrete situation with a patient. A repetition of the conducted study with experienced physicians working in their field for years could be an interesting way to further gain inside into the prioritization procedure.

Although willingness to pay is a measure often difficult to define and work with, the presented investigation demonstrates that physicians take it into account for their prioritization decision. The conducted experiment is specific as the elicitation of individual willingness to pay is seldom realized that sophisticated:

- The first particular feature is that the maximum willingness to pay was restricted and could not exceed the amount of five Euros. In case that a subject indicated a willingness to pay of five Euros, the experimenter reduced the overall payment each individual received for the participation in the experiment, so that the subject received not 15 but 10 Euros. Thus, effects coming from general income could be circumvented. Additionally, the sample was made up of students only, which means that within this population, the ability to pay higher amounts is not very high anyway.

- The second special feature concerning willingness to pay in the experiment is the way how the individual maximum willingness to pay was elicited. The elicitation procedure to find the individual willingness to pay of each of the five subjects took into account that subjects might behave strategically and generally indicated the maximum willingness to pay of five Euros to influence the decision maker. Other before-mentioned studies could demonstrate that the ability of patients to complain or to foster treatment have a significant effect on the physicians decision to realize specific therapies or to treat this individual first (Strech et al. 2008b). As a consequence, the presented experiment elicits willingness to pay within an extra step of the experiment. Subjects learn that there are two situations in which it is possible in the course of the experiment that they have to immerse their hand in the allocated water bowl with the individual temperature and immersion duration: an immersion results in case that the decision maker did not set them free, and a second immersion is possible in the end, when one individual willingness to pay decisions is realized for a randomly defined subject. As the instructions were explained in detail, the patients have to indicate at the beginning which amount of money they are willing to pay as a maximum to avoid the pain. This had also been illustrated in table 1. Eleven choices had to be made in which the pain is always contrasted with rising amounts of money. In the end, one out of the five patients was determined randomly. For her, the eleven choices were displayed and again one of these was determined randomly. In case that the cold pressor pain had been selected, it was realized immediately, otherwise, the amount of money in the determined choice was deducted from the payment for participation. This procedure makes sure that no individual indicates a willingness which was only selected for strategic and thus manipulative reasons. Each participant knew that their willingness to pay indication could end up in a loss of one third (in case of the maximum willingness to pay) of the overall payment.

- The third specific aspect represented by the conditions how willingness to pay was used in this experiment refers to the equal situation of the five subjects. They all had the same amount of money at their disposal: five Euros maximum willingness to pay which was deduced from their upcoming earnings for participation. Accordingly, their current private wealth was not directly relevant for the individual to make his decision. Thereby, the influence of the background each subjects has is strongly reduced. This also has a positive influence on the decider who did not come into any considerations on the fairness of the willingness to pay parameter maybe reflecting the individual financial situations. This can be different where real willingness to pay from real patients suffering from an illness is questioned. Here it can easily have an unethical connotation if treatment is given to those only paying high prices for it, especially in health care systems financed on the basis of solidarity. This can hardly be realized and of course is not the aim of this research. Instead, the concrete elicitation process of real willingness to pay helps to understand with what kind of information physicians are able and willing to work.

As explained above, the experimental situation was constructed with the goal to mirror the high level of complexity which is the everyday challenge of a physician. When additional allocation decisions are expected to be handled by these physicians at the same time, their decision making becomes even more difficult. Accordingly, it is interesting to analyze if there can be situations and procedures that end up in a parameter that contains other parameters so that complexity is reduced. This is what the QALY concept tries to deliver. But neither approach QALY nor willingness to pay so far is free from criticism and is applicable for every situation. This is why it must be evaluated for different questions which method helps best to come to the relevant results and conclusions. This also refers to the differentiation of Cost-Benefit-Analysis versus Cost-Utility-Analysis or the question if all relevant parameters are available for the implementation in a study or if there are aspects that cannot be included for whatever reason. Chapter 1 demonstrates one scenario where willingness to pay can be evaluated as a helpful parameter and in which a specific way to elicit it is realized. Here, it can be interpreted as a kind of content validity accepted by the expert to base decisions on. Further analysis is necessary to define appropriate contexts and questions where willingness to pay has resilient advantages over QALYs or other approaches and where it can be implemented to reduce high levels of complexity.

In chapter 2, the focus lays on the preference and decision behavior individuals show over two negative alternatives, i. e. two painful experiences. Randomized over the sample, each subject went through two sequences of pain induction. These sequences varied only concerning the ordering of pain intensities, either increasing pain or declining pain. This scenario is comparable to the situation of a patient for whom two treatment options are optional. After the patient has tried both of them, she might be in the situation to decide with which of the two she wants to continue. Independent from the treatment success, it can be assumed that both treatments differ in the levels of well-being during the procedure or in other parameters which can vary over the time of the treatment. The conducted study controls for equal parameters within the two treatments but makes sure that they do not occur in the same temporal ordering. Only the ordering of equal elements over time might change the overall evaluation of the two experiences.

The second experiment demonstrates in particular the importance of real consequences, especially in cases where the research question refers to the differentiation of rational from irrational behavior. An investigation of preferences with rating scales or a simple preference request would have delivered the result that individual behavior does not follow predictions from solid frameworks such as Prospect Theory. It rather seems to support the contrary, namely that a broad bundle of different behavioral anomalies are necessary to forecast how subjects behave in different situations. The general preference indication in the sample remains irrational or anomaly-conform before real consequences from preferences are included in terms of monetary disadvantages.

Moreover, the described second experiment and its results emphasize the relevance of an extensive data analysis. The main effect within the sample depends on the strict randomization of the first and the second experience, as the former operates as a reference point for the latter. As a consequence, the interpretation of the data must take thoroughly into account which sequence was the first one experienced by each individual. As the application of the data to Cumulative Prospect Theory demonstrates, the reference point assumption in combination with the development of the experience as better or worse clearly comes to the result that the behavioral pattern occurring within the sample fits perfect with the theory. This study thus also aims to emphasize the importance to randomize existing options within experiments, especially concerning the first sequence of every procedure in the laboratory.

The similarity between the first two chapters refers to the usage of willingness to pay or rather real willingness to pay. Although the specific context for which this method was implemented differs, the striking advantage is central in both studies: willingness to pay needs to be for real. For the first study, this means that the deciders can rely on the information as a mirror for individual urgency. In the second study, the preference elicitation with willingness to pay makes subjects reflect in more detail about their preferences so that a former indicated favoring of one sequences equalizes and the preference structure transforms into indifference.

The third investigated problem in chapter 3 refers to risk preferences over two different outcomes: pain intensity and pain duration. The risk attitudes over these two outcomes both refer to different contexts.
Analogous with the second experiment, one focus lay on the application of Prospect Theory and its predictions to a health-related context. Working with pain induction, the experience subjects made in the conducted experiments can be framed as losses. Prospect Theory says that individuals are risk seeking for losses and risk averse for gains. According to that, subjects should be risk seeking both for pain intensity and pain duration.

The second construct in chapter 3 which should be investigated in more detail was the QALY concept. One aspect of this approach was defined, namely the assumption of risk neutrality over time. It is hard to believe that individuals are risk neutral when it comes to their own life time. Coming from Prospect Theory, one would assume that individuals are risk seeking when the scenario is framed as a loss,

whereas they behave risk averse for scenarios where life years can be gained. Accordingly, the conducted study aimed at investigating if risk neutrality is a justified assumption.
The cold pressor test formed the basis to operationalize pain in intensity and duration. By means of a Holt and Laury procedure, individual risk attitudes were elicited for pain intensity and pain duration. The results raise criticism over both, assumptions of Prospect Theory and the QALY concept. Subjects are risk averse over pain intensity and risk seeking over pain duration. Accordingly, Prospect Theory predicts the behavior correct for losses in terms of pain intensity, but not for losses in terms of duration. Risk neutrality was found in neither scenario. The QALY concept thus represents quality of life based on a false assumption.

It is recommended to further investigate individual risk attitudes in terms of life time so that the concept and the calculation of the QALY index can be adapted accordingly. A special emphasis could lay on the repetition of appropriate experiments in this context which also include real consequences from decision making, so that any hypothetical bias in this already complex decision situation can be prevented.

The consequence of the presented research and the described results is a recommendation especially to health economists to enlarge working with experiments including real consequences from decision making or preference statements. It is obvious that experiments and the results they generate are limited. However, the more involving character of settings with real consequences from decision making can contribute to and enlarge the existing findings within a specific topic. It definitely makes a difference whether the individual subject engaged in the experiment is aware that his current behavior and the decision he takes right now will result in a painful consequences in the next step. Thereby, the character of the consequence is not essential; it might be pain, a monetary loss or any other kind of loss or gain. There are still many topics where we are far from precisely predicting preferences and behavior, and we should not let this relevant factor out on our way towards a deeper understanding.

Different fundamental questions have been part of the presented investigations. Coming from a health-system-related scenario where resource allocation and prioritization are current difficulties as well as fairness and rationality within such complex tasks, this dissertation continued with the application of Prospect Theory. It demonstrates that Cumulative Prospect Theory instead of well known behavioral anomalies is of high explanatory power when it comes to real decision making. Finally, this dissertation investigates questions concerning the measurement of well-being with the QALY concept and willingness to pay.

This work adds new findings to existing results and did its stint to further optimize research methods and maybe more important, the understanding of the individual.

Bibliography

Abdellaoui, M., Bleichrodt, H., & L`Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, *36*, 245–266.

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science*, *53*, 1659–1674.

Acton, J. P. (1976). *Evaluating Public Programs to Save Lives: the Case of Heart Attacks.* (Report No. R950RC). Santa Monica: RAND Corporation.

Adams, A., Soumerai, S., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care*, *11*, 187–192.

Andersen, S., Harrison, G., & Rutström, E. (2006). Choice Behavior, Asset Integration and Natural Reference Points. *Working Paper, 06-07*.

Ardal, S., Butler, J., Edwards, R., & Lawrie, L. (2008). *Priority Setting: The Health Planner`s Toolkit*.

Ariely, D. (1998). Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *Journal of Behavioral Decision Making*, *11*, 19–45.

Ariely, D., Kahneman, D., & Loewenstein, G. (2000). Joint comment on "When does duration matter in judgment and decision making?". *Journal of Experimental Psychology: General*, *129*, 524–529.

Ariely, D., & Loewenstein, G. (2000). When does duration matter in judgement and decision making? *Journal of Experimental Psychology: General*, *129*, 508–523.

Atkinson, R., & Shiffrin, R. (1968). Human memory: a proposed system and its control processes. In K. Spence, & J. Spence (Eds.), *The psychology of learning and motivation: advances in research and theory* (pp. 90–191). New York: Academic Press.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove [England] ;, New York: Psychology Press.

Barry, M. J., & Edgman-Levitan, S. (2012). Shared Decision Making — The Pinnacle of Patient-Centered Care. *New England Journal of Medicine*, *366*, 780–781.

Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2008). Three studies on the neuroeconomics of decision-making when payoffs are real and negative. *Advances in Health Economics and Health Services Research*, *20*, 1–29.

Birnie, K. A., Noel, M., Chambers, C. T., Baeyer, C. L. von, & Fernandez, C. V. (2011). The Cold Pressor Task: Is it an Ethically Acceptable Pain Research Method in Children? *Journal of Pediatric Psychology*, *36*, 1071–1081.

Bleichrodt, H., Wakker, P., & Johannesson, M. (1997). Characterizing QALYs by Risk Neutrality. *Journal of Risk and Uncertainty*, *15*, 107–114.

Blumenschein, K., Johannesson, M., Yokoyama, K. K., & Freeman, P. R. (2001). Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *Journal of Health Economics*, *20*, 441–457.

Brown, L. D. (1991). The national politics of Oregon's rationing plan. *Health Affairs*, *10*, 28–51.

Buxton, M. J., & Chambers, J. D. (2011). What values do the public want their haelth care system to use in evaluating technologies? *European Journal of Health Economics*, *12*, 285–288.

Chéry-Croze, S. (1983). Relationship between noxious cold stimuli and the magnitude of pain sensation in man. *Pain*, *15*, 265–269.

Chuck, A., Adamowicz, W., Jacobs, P., Ohinmaa, A., Dick, B., & Rashiq, S. (2009). The Willingness to Pay for Reducing Pain and Pain-Related Disability. *Value in Health*, *12*, 498–506.

Conner-Spady, B., & Suarez-Almazor, M. (2003). Variation in the Estimation of Quality-adjusted Life-years by Different Preference-based Instruments. *Medical Care*, *41*, 791–801.

Cookson, R. (2003). Willingness to pay methods in health care: a sceptical view. *Health Economics*, *12*, 891–894.

Davis, R. (1963). Recreation planning as an economic problem. *Natural Resources Journal*, *3*, 239–249.

Deese, J., & Kaufman, R. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, *54*, 180–187.

Demaree, H., Shenal, B., Everhart, E., & Robinson, J. (2004). Primacy and Recency Effects Found Using Affective Word Lists. *Cognitive and Behavioral Neurology*, *17*, 102–108.

Diederich, A., Swait, J., Wirsik, N., & Dowdy, D. W. (2012). Citizen Participation in Patient Prioritization Policy Decisions: An Empirical and Experimental Study on Patients' Characteristics. *PLoS ONE*, *7*, e36824.

Diener, E., Wirtz, D., & Oishi, S. (2001). End Effects of Rated Life Quality: The James Dean Effect. *Psychological Science*, *12*, 124–128.

Do, A. M., Rupert, A. V., & Wolford, G. (2008). Evaluations of pleasurable experiences: The peak-end rule. *Psychonomic Bulletin & Review*, *15*, 96–98.

Dolan, P., Richard, E., & Aki, T. (2008). *Microsoft Word - Final report revised.doc: final report*.

Donaldson, C., & Shackley, P. (2003). Wilingness to Pay for Health Care. In A. Scott, A. Maynard, & R. F. Elliott (Eds.), *Advances in health economics* (pp. 1–24). New York: J. Wiley.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O´Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes.* (3rd). Oxford ;, New York: Oxford University Press.

Finn, B. (2010). Ending on a high note: Adding a better end to effortful study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1548–1553.

Fredrickson, B. L., & Kahneman, D. (1993). Duration Neglect in Retrospective Evaluations of Affective Episodes. *Journal of Personality and Social Psychology*, *65*, 45–55.

Fuchs, C. (2010). Priorisierung aus Sicht der Medizin - Positionsbestimmung der Ärzteschft zum Thema. In H. Lohmann, & U. Preusker (Eds.), *Priorisierung statt Rationierung: Zukunftssicherung für das Gesundheitssystem* (pp. 15–32). Heidelberg ;, München, Landsberg, Frechen, Hamburg: Economica-Verl.

Glenberg, A., Bradley, M., Stevenson, J., & Kraus, T. (1980). A Two-Process Account of Long-Term Serial Position Effects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 355–369.

Godolphin, W. (2009). Shared Decision-Making. *Healthcare Quarterly*, *12*, 186–190.

Green, C. (2009). Investigating public preferences on 'severity of health' as a relevant condition for setting healthcare priorities. *Social Science and Medicine*, *68*, 2247–2255.

Greiner, B. (2004). The Online Recruitment System ORSEE – A Guide for the Organization of Experimens in Economics. In K. Kremer, & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen* (pp. 79–93). Göttingen: Gesellschaft für wissenschaftliche Datenverarbeitung.

Ham, C. (1997). Priority setting in health care: learning from international experience. *Health Policy*, *42*, 49–66.

Harrison, G., & Rutström, E. (2009). Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral. *Experimental Economics*, *12*, 133–158.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Physiological Reviews*, *93*, 258–268.

Hauerwaas, A. (2008). Die Rolle von Innovationen am Forschungsstandort Deutschland in Bezug auf das Gesundheitswesen: *Innovationen im Gesundheitswesen.* (1st ed.). s.l: GRIN Verlag.

Hines, E. A., & Brown, G. E. (1936). The cold pressor test for measuring the reactibility of the blood pressure: data concerning 571 normal and hypertensive subjects. *The American Heart Journal*, *11*, 1–9.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, *92*, 1644–1655.

Holt, C. A., & Laury, S. K. (2005). Risk Aversion and Incentive Effects: New Data without Order Effects. *American Economic Review*, *95*, 902–904.

Jones-Lee, M. (1974). The value of changes in the probability of death or injury. *Journal of Political Economy*, *82*.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*, 401–405.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*, 263–291.

Keogh, E., Hatton, K., & Elery, D. (2000). Avoidance and focused attention and the perception of pain: differential effects for men and women. *Pain*, *85*, 225–230.

Keogh, E., & Herdenfeldt, M. (2002). Gender, coping and the perception of pain. *Pain*, *97*, 195–201.

Klein, R. (1993). Dimensions of rationing: who should do what? *BMJ*, *307*, 309–311.

Kroll, E., Trarbach, J. & Vogt, B. (2011). Determining risk preferences for pain. (Working Paper). Faculty of Management and Economics, Otto-von-Guericke University Magdeburg.

Kroll, E.; Trarbach, J. & Vogt, B. (2012). Do people have a preference for increasing or decreasing pain? – An experimental comparison of psychological and economic measures in health-related decision making (Working Paper). Faculty of Management and Economics, Otto-von-Guericke University Magdeburg.

Li, Y., & Epley, N. (2009). When the best appears to be saved for last: Serial position effects on choice. *Journal of Behavioral Decision Making*, *22*, 378–389.

Lorenz, J. (2002). Quantifizierung des Schmerzes. In H. Beck, E. Martin, J. Motsch, & J. Schulta am Esch (Eds.), *Schmerztherapie: Intensivmedizin, Notfallmedizin, Schmerztherapie* (pp. 51–58). Stuttgart ;, New York: Georg Thieme.

Lovallo, W. (1975). The Cold Pressor Test and Autonomic Function: A Review and Integration. *Psychophysiology*, *12*, 268–282.

Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in Choice: Effects of Serial Position on Preferences. *Psychological Science*, *20*, 1309–1312.

Marckmann, G. (2009). Priorisierung im Gesundheitswesen: Was können wir aus den internationalen Erfahrungen lernen? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *103*, 85–91.

Martijn Willemsen & Eric Johnson (2006). *Mouse Lab*: Decision Science News.

McNeil, B. J., Weichselbaum, R., & Pauker, S. G. (1978). Fallacy of the five-year survival in lung cancer. *New England Journal of Medicine*, *299*, 1379–1401.

McNeil, B. J., Weichselbaum, R., & Pauker, S. G. (1981). Speech and Survival — Tradeoffs between Quality and Quantity of Life in Laryngeal Cancer. *New England Journal of Medicine*, *305*, 982–987.

Mitchell, L. A., MacDonald, R. A., & Brodie, E. E. (2004). Temperature and the cold pressor test. *The Journal of Pain*, *5*, 233–237.

Mitton, C., & Donaldson, C. (2003). Tools of the trade: a comparative analysis of approaches to priority setting in healthcare. *Health Services Management Research*, *16*, 96–105.

Miyamoto, J. M., & Eraker, S. A. (1985). Parameter Estimates for a QALY Utility Model. *Medical Decision Making*, *5*, 191–213.

Morgenstern, R., Heldmann, M. & Vogt, B. (2013a). Differences in cognitive control between real and hypothetical payoffs. (Working Paper). Faculty of Management and Economics, Otto-von-Guericke University Magdeburg.

NICE (2008). *SOCIAL VALUE JUDGEMENTS Principles for the development of NICE guidance.* (2nd ed.).

O'Brien, B. J., Spath, M., Blackhouse, G., Severens, J., Dorian, P., & Brazier, J. (2003). A view from the bridge: agreement between the SF-6D utility algorithm and the Health Utilities Index. *Health Economics*, *12*, 975–981.

Oliver, A., & Cookson, R. (2010). Analysing risk attitudes to time. *Health Economics*, *19*, 644–655.

Parker, R. (1975). Social administration and scarcity. In E. Butterworth, & R. Holman (Eds.), *Social welfare in modern Britain* (pp. 204–212). [London]: Fontana.

Pesheva, D., Kroll, E. B., & Vogt, B. (2011). Gender differences in willingness to pay to avoid pain and their correlation with risk. *Journal of Neuroscience, Psychology, and Economics*, *4*, 181–191.

Pliskin, J. S., Shepard, D. S., & Weinstein, M. C. (1980). Utility Functions for Life Years and Health Status. *Operations Research*, *28*, 206–224.

Ratcliffe, J., Buxton, M., Young, T., & Longworth, L. (2005). Determining Priority for Liver Transplantation: A Comparison of Cost per QALY and Discrete Choice Experiment-Generated Public Preferences. *Applied Health Economics and Health Policy*, *4*, 249–255.

Redelmeier, D. A., & Kahneman, D. (1996). Patient´s memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*, 3–8.

Redelmeier, D. A., Katz, J., & Kahneman, D. (2003). Memories of colonoscopy: a randomized trial. *Pain*, *104*, 187–194.

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12*, 179–195.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1446–1465.

Rosen, A. B., Tsai, J. S., & Downs, S. M. (2003). Variations in Risk Attitude across Race, Gender, and Education. *Medical Decision Making*, *23*, 511–517.

Ryynänen, O.-P., Markku, M., Kinnunen, J., & Takala, J. (1999). Attitudes to health care prioritisation methods and criteria among nurses, doctors, politicians and the general public. *Social Science and Medicine*, *49*, 1529–1539.

Ryynänen, O.-P., Myllykangas, M., Kinnunen, J., Halonen, P., & Takala, J. (2000). PRIORITIZATION ATTITUDES AMONG DOCTORS AND NURSES EXAMINED BY A SCENARIO METHOD. *International Journal of Technology Assessment in Health Care*, *16*, 92–99.

Sabik, L. M., & Lie, R. K. (2008). Priority setting in health care: Lessons from the experiences of eight countries. *International Journal for Equity in Health*, *7*, 4.

Schattner, A., Rudin, D., & Jellin, N. (2004). Good physicians from the perspective of their patients. *BMC Health Services Research*, *4*, 26.

Scholl, M., Segmüller, T., Zegelin, A., Wagner, F., & Bienstein, C. (2013). *Menschen würdig pflegen? Das Recht auf qualifizierte Pflege: Eine Diskussionsschrift*.

Schreiber, C. A., & Kahneman, D. (2000). Determinants of Remembered Utility of Aversive Sounds. *Journal of Experimental Psychology: General*, *129*, 27–42.

Schwappach, D. L. (2003). Does it matter who you are or what you gain? an experimental study of preferences for resource allocation. *Health Economics*, *12*, 255–267.

Schultheiss, C. (2004). Im Räderwerk impliziter Rationierung. *Psychoneuro*, 30, 568-574

Stacey, D., Bennett, C. L., Barry, M. J., Col, N. F., Eden, K. B., Holmes-Rovner, M., Llewellyn-Thomas, H., Lyddiatt, A., Légaré, F., & Thomson, R. (2011). Decision aids for people facing health treatment or screening decisions. *Cochrane Database Systematic Review, 5*.

Stiggelbout, A., Kiebert, G., Kievit, J., Leer, J., Stoter, G., & Haes, J. de (1994). Utility Assessment in Cancer Patients: Adjustment of Time Tradeoff Scores for the Utility of Life Years and Comparison with Standard Gamble Scores. *Medical Decision Making*, *14*, 82–90.

Strech*, D., Börchers*, K., Freyer*, D., Neumann*, A., Wasem*, J., & Marckmann*, G. (2008a). Ärztliches Handeln bei Mittelknappheit. *Ethik in der Medizin*, *20*, 94–109.

Strech, D., Synofzik, M., & Marckmann, G. (2008b). How Physicians Allocate Scarce Resources at the Bedside: A Systematic Review of Qualitative Studies. *Journal of Medicine and Philosophy*, *33*, 80–99.

Streff, A., Kuehl, L. K., Michaux, G., & Anton, F. (2010). Differential physiological effects during tonic painful hand immersion tests using hot and ice water. *European Journal of Pain*, *14*, 266–272.

Thaler, R. H. & Johnson, E. J. (1990). Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science, 36*, 643-660

Tu, Q. (2005). *Empirical Analysis of Time Preferences and Risk Aversion.* CentER PhD Thesis 142, 142. Tilburg: Tilburg University.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Unützer, J., Katon, W. J., Russo, J., Simon, G., & Von Koeff, M. (2003). Willingness to pay for depression treatment in primary care. *Psychiatric Services*, *54*, 340–345.

van de Mortel, T. (2008). Faking it: social desirability response bias in selfreport research. *AUSTRALIAN JOURNAL OF ADVANCED NURSING*, *25*.

van Delden, J. J. M. (2004). Medical decision making in scarcity situations. *Journal of Medical Ethics*, *30*, 207–211.

Varey, C., & Kahneman, D. (1992). Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making*, *5*, 169–185.

Verhoef, L. C., Haan, A. F. de, & van Daal, W. A. (1994). Risk Attitude in Gambles with Years of Life: Empirical Support for Prospect Theory. *Medical Decision Making*, *14*, 194–200.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. (1st ed.). Princeton: Princeton University Press.

Weinstein, M. C., & Stason, W. B. (1977). Foundations of Cost-Effectiveness Analysis for Health and Medical Practices. *New England Journal of Medicine*, *296*, 716–721.

Weinstein, M. C., Torrance, G., & McGuire, A. (2009). QALYs: The Basics. *Value in Health*, *12*, S5–S9.

Wernitz, M. H., & Pelz, J. (2011). *Gesundheitsökonomie und das deutsche Gesundheitswesen: Ein praxisorientiertes Lehrbuch für Studium und Beruf.* (1st ed.). Stuttgart: Kohlhammer.

Wodarg, W. (2010). Priorisierung ja, aber transparent und demokratisch. Eine Kritik der impliziten Priorisierung. In H. Lohmann, & U. Preusker (Eds.), *Priorisierung statt Rationierung: Zukunftssicherung für das Gesundheitssystem* (pp. 93–111). Heidelberg ;, München, Landsberg, Frechen, Hamburg: Economica-Verlag.

Zentrale Ethikkommission (2007). Priorisierung medizinischer Leistungen im System der Gesetzlichen Krankenversicherung (GKV). *Deutsches Ärzteblatt*, *104*, 2750–2754.

Appendix A

Does your physician see your needs? An experimental analysis on prioritizing patients

Written experimental instructions

Einverständniserklärung

Ich bin aufgeklärt worden, dass es sich bei diesem Experiment um ein Schmerzexperiment mit kaltem Wasser handelt. Es besteht dabei kein Risiko für meine Gesundheit. Wenn es nötig ist, kann ich das Experiment zu jeder Zeit abbrechen.

Datum Unterschrift

Anleitung

Herzlich willkommen zu diesem Experiment und vielen Dank für Ihre Teilnahme. Sie erhalten 10,00 Euro für die Teilnahme an diesem Experiment. Bitte lesen Sie diese Anweisungen – die für alle gleich sind – sorgfältig durch.

Bitte verhalten Sie sich ruhig und schalten Sie Ihre Mobiltelefone aus. Kommunikation zwischen den Teilnehmern ist nicht erlaubt. Wenn Sie eine Frage haben, heben Sie bitte die Hand. Ein Experimentleiter wird dann zu Ihnen kommen und Ihre Frage unter vier Augen beantworten. Leisten Sie diesen Regeln nicht Folge, müssen wir Sie vom Experiment ausschließen. Für diesen Fall werden Sie von sämtlichen Zahlungen für Ihre Teilnahme ausgeschlossen.

Ihre Auszahlung werden wir am Ende des Experiments mit Ihnen abrechnen.

Rollenzuordnung

Zu Beginn des Experiments werden alle Teilnehmer einer von zwei Rollen (Typ A oder Typ B) zugeordnet. Dabei wird ein Teilnehmer Typ A und fünf Teilnehmer Typ B bestimmt. Abhängig von Ihrer Rolle stehen Sie unterschiedlichen Entscheidungssituationen gegenüber. Im Folgenden wird immer explizit dargestellt, welche Entscheidungen von welchen Rollen getroffen werden.

Alle Teilnehmer in diesem Experiment mit Typ B befinden sich in einem Raum. Der Teilnehmer mit Typ A befindet sich getrennt von den Teilnehmern mit Typ B in einem anderen Raum auf dem Campus der Universität Magdeburg. Neben seinem Typ erhält jeder Teilnehmer zu Beginn des Experiments eine Identifikationsnummer, über die er eindeutig bestimmt werden kann.

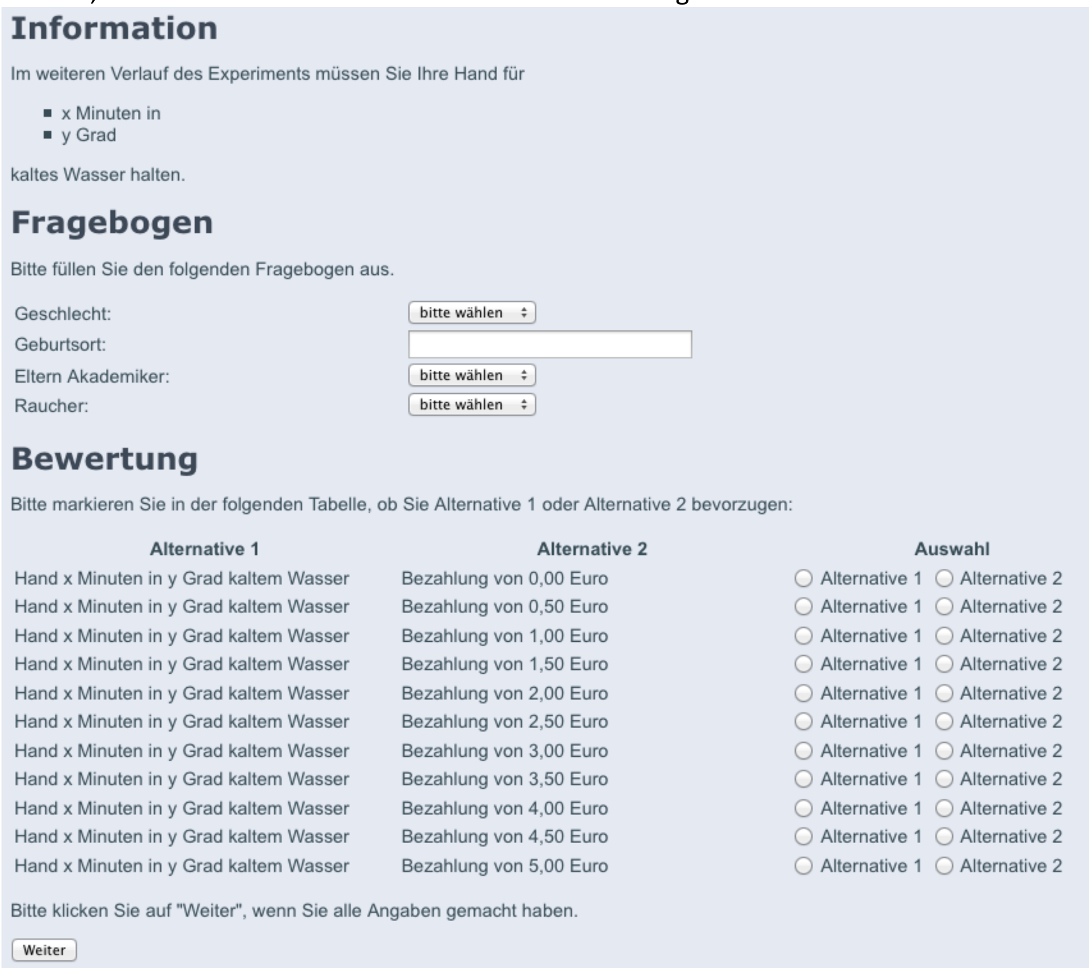
<u>Ihre Rolle ist Typ A</u>

Ablauf des Experiments

Dieses Experiment besteht aus vier Stufen. In der ersten Stufe machen alle Teilnehmer, egal ob sie vom Typ A oder Typ B sind, eine Schmerzerfahrung. In der zweiten Stufe erhält der Teilnehmer von Typ A die Möglichkeit, einem Teil der Teilnehmer von Typ B weitere Schmerzen zu ersparen. In der dritten Stufe beantworten alle Teilnehmer zwei Fragebögen zu ihren Entscheidungen. Sie erhalten hierfür einen getrennten Fragebogen, den Sie unabhängig von den anderen Teilnehmern beantworten. In der vierten Stufe erleiden die Teilnehmer von Typ B einen Schmerz, wenn der Teilnehmer von Typ A sie nicht von diesem Schmerz befreite.

Stufe 1: Schmerzerfahrung

In jedem der beiden Räume befindet sich ein Bassin mit 7° Celsius Grad kaltem Wasser. In dieser Stufe müssen alle Teilnehmer (von Typ A und Typ B) nacheinander ihre Hand für exakt 20 Sekunden in dieses Bassin halten. Anschließend wird den Teilnehmern von Typ B die Dauer und die Temperatur mitgeteilt, bei der sie im weiteren Verlauf des Experiments ihre Hand in Wasser eintauchen müssen. Diese liegt zwischen 4° und 12° Celsius. Danach erhalten die Teilnehmer von Typ B einen Fragebogen. In diesem Fragebogen machen die Teilnehmer von Typ B einige persönliche Angaben (siehe Abb. 1). Zusätzlich wird den Teilnehmern von Typ B eine Tabelle mit 11 Zeilen angezeigt. In jeder Zeile stehen zwei Alternativen. Alternative 1 beschreibt eine Schmerzerfahrung, während Alternative 2 einen Verlust zwischen 0 Euro und 5 Euro beschreibt. Die Teilnehmer von Typ B geben für jede Zeile an, ob sie Alternativen 1 oder Alternative 2 bevorzugen.



Information

Im weiteren Verlauf des Experiments müssen Sie Ihre Hand für

- x Minuten in
- y Grad

kaltes Wasser halten.

Fragebogen

Bitte füllen Sie den folgenden Fragebogen aus.

Geschlecht:	bitte wählen
Geburtsort:	
Eltern Akademiker:	bitte wählen
Raucher:	bitte wählen

Bewertung

Bitte markieren Sie in der folgenden Tabelle, ob Sie Alternative 1 oder Alternative 2 bevorzugen:

Alternative 1	Alternative 2	Auswahl
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 0,00 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 0,50 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 1,00 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 1,50 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 2,00 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 2,50 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 3,00 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 3,50 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 4,00 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 4,50 Euro	◯ Alternative 1 ◯ Alternative 2
Hand x Minuten in y Grad kaltem Wasser	Bezahlung von 5,00 Euro	◯ Alternative 1 ◯ Alternative 2

Bitte klicken Sie auf "Weiter", wenn Sie alle Angaben gemacht haben.

Weiter

Abb 1. Fragebogen für Typ B

Stufe 2: Entscheidung über weitere Schmerzen

Zu Beginn dieser Stufe erhält der Teilnehmer von Typ A ein Budget in Höhe von 5,00 Euro. Der Teilnehmer von Typ A kann diesen Betrag nutzen, um einem Teil der Teilnehmer von Typ B weitere Schmerzen zu ersparen. Dafür sieht der Teilnehmer von Typ A auf seinem Monitor eine Liste aller Teilnehmer von Typ B mit ihrer jeweiligen Identifikationsnummer (siehe Abb. 2). In dieser Liste befinden sich neben der Identifikationsnummer folgende weitere Informationen über jeden Teilnehmer von Typ B:

- Geschlecht
- Temperatur beim Handeintauchen
- Dauer des Handeintauchens
- Geburtsort
- Eltern Akademiker
- Betrag in Euro den Teilnehmer in Stufe 1 angab, um Hand nicht erneut in Wasser halten zu müssen
- Kosten in Euro für Teilnehmer von Typ A
- Raucher

Diese Informationen sind für den Teilnehmer von Typ A nicht gleichzeitig sichtbar, sondern sie können einzeln nacheinander, aber in beliebiger Reihenfolge und Wiederholung gelesen werden. Dazu bewegt der Teilnehmer von Typ A den Mauszeiger über das entsprechende Feld auf seinem Computerbildschirm. Nachdem der Mauszeiger dieses Feld wieder verlassen hat, verschwindet die Information wieder.

Der Teilnehmer von Typ A kann diese Informationen nutzen, um einen Teil der Teilnehmer von Typ B vom Handeintauchen freizustellen. Der Teilnehmer von Typ A kann maximal so viele Teilnehmer vom Handeintauchen freistellen, wie sein Budget umfasst. Schöpft der Teilnehmer von Typ A sein Budget nicht aus, verfällt der verbleibende Betrag.

Beispiel: Der Teilnehmer von Typ A entscheidet sich, 2 Teilnehmern weitere Schmerzen zu ersparen. Die Kosten in Euro für Teilnehmer von Typ A für den Ersten sind 2,50 Euro, die für den Zweiten sind 2,30 Euro. Damit entstehen ihm Kosten in Höhe von 4,80 Euro. Da sein Budget 5,00 Euro beträgt, verfällt der verbleibende Betrag in Höhe von 0,20 Euro.

Nachdem der Teilnehmer von Typ B diese Entscheidung für die Teilnehmer von Typ A, die zeitgleich an diesem Experiment teilnehmen, getroffen hat, entscheidet er für zwei weitere Gruppen von Teilnehmern von Typ B. Hierfür werden ihm analoge Informationen auf dem Bildschirm angezeigt. Diese zusätzlichen Gruppen von Teilnehmern von Typ B werden im Nachgang an dieses Experiment ausbezahlt.

Abb 2. Entscheidungsbildschirm Typ A

Stufe 3: Fragebögen

Nachdem der Teilnehmer von Typ A seine Entscheidung getroffen hat, erhalten sowohl die Teilnehmer von Typ A als auch von Typ B zwei weitere Fragebögen.

Stufe 4: Schmerz für Teilnehmer von Typ B

Im Anschluss an die Fragebögen wird die Entscheidung aus Stufe 2 in dieser Stufe umgesetzt. D.h. dass alle Teilnehmer von Typ B erfahren, ob sie vom Typ A-Spieler freigestellt wurden und ihre Hand nicht in Wasser eintauchen müssen.

Anschließend müssen die Teilnehmer ihre Hand gemäß der gewählten Entscheidung in Wasser eintauchen. Die Teilnehmer von Typ B tauchen hierfür ihre Hand gleichzeitig in das Wasser ein. D.h. die Teilnehmer, die ihre Hand kürzer in Wasser eintauchen müssen, beenden diese Stufe früher.

Der Teilnehmer von Typ A verfolgt diesen Vorgang auf seinem Monitor mit Hilfe einer Webcam. Damit sichergestellt ist, dass die Teilnehmer vom Typ B tatsächlich in der aktuellen Sitzung teilnehmen, befindet sich im Bild ein Exemplar der Magdeburger Volksstimme des aktuellen Tages und eine Uhr mit der aktuellen Uhrzeit. Sobald alle Teilnehmer von Typ B ihre Hand wieder aus dem Wasser genommen haben, folgt die Auszahlung.

Abschließende Auszahlung

Alle Teilnehmer von Typ A und von Typ B erhalten 10,00 Euro für die Teilnahme am Experiment. Zusätzlich wird für einen der Teilnehmer von Typ B eine seiner Angaben im Fragebogen (siehe Abbildung 1, Abschnitt Bewertung) von Schritt 1 real. Hierfür wird zunächst der entsprechende Teilnehmer ausgelost. Dafür werden Kugeln mit den Zahlen 1 bis 5 in eine Urne gelegt. Anschließend wird genau eine der Kugeln verdeckt gezogen. Die Realisierung gilt für den Teilnehmer, dessen Identifikationsnummer mit der gezogenen Nummer übereinstimmt. Dafür wird eine zweite Urne mit den Zahlen von 1 bis 11 gefüllt, und eine der Kugeln verdeckt gezogen. Ist die gezogene Kugel die Kugel mit der 1, so wird die erste Zeile realisiert, steht auf ihr die 2, so wird die zweite Zeile realisiert usw. bis zur elften Zeile. Bei der Umsetzung wird im ausgefüllten Fragebogen abgeglichen, ob der Teilnehmer von Typ B in Schritt 1 Alternative 1 oder Alternative 2 bevorzugte. Bevorzugte er Alternative 1, so muss er seine Hand ein zweites Mal in kaltes Wasser halten, bevorzugte er Alternative 2, so wird seine Auszahlung von 10,00 Euro um den Betrag in der ausgelosten Zeile reduziert.

Fragebogen

(1) Bitte beschreiben Sie, wie Sie Ihre Entscheidung getroffen haben. Beschreiben Sie Ihre Strategie dabei so ausführlich, dass damit ein Computerprogramm geschrieben werden könnte, dass Ihre Strategie umsetzt.

(2) Bitte markieren Sie, welche Eigenschaften der Teilnehmer von Typ B Sie wie stark beeinflusst haben (0 = Keine Beeinflussung, 10 = sehr starke Beeinflussung).

Eigenschaft	**Grad der Beeinflussung**										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Geschlecht des Teiln. von Typ B											
Temperatur beim Handeintauchen											
Dauer des Handeintauchens											
Geburtsort des Teiln. von Typ B											
Eltern von Teiln. mit Typ B sind Akademiker											
Betrag in Euro den Teiln. B in Stufe 1 angab											
Kosten in Euro für Teiln. von Typ A											
Teiln. von Typ B ist Raucher											

Mit Hilfe der von Ihnen in diesem Fragebogen beschriebenen Strategie wird ein weiteres Experiment durchgeführt.

Fragebogen

(1) Bitte beschreiben Sie, wie Sie glauben, dass der Spieler von Typ A seine Entscheidung getroffen hat. Beschreiben Sie seine Strategie dabei so ausführlich, dass damit ein Computerprogramm geschrieben werden könnte, dass seine Strategie umsetzt.

(2) Bitte markieren Sie, welche Eigenschaften der Teilnehmer von Typ B den Teilnehmer von Typ A wie stark beeinflusst haben (0 = Keine Beeinflussung, 10 = sehr starke Beeinflussung).

Eigenschaft	**Grad der Beeinflussung**										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Geschlecht des Teiln. von Typ B											
Temperatur beim Handeintauchen											
Dauer des Handeintauchens											
Geburtsort des Teiln. von Typ B											
Eltern von Teiln. mit Typ B sind Akademiker											
Betrag in Euro den Teiln. B in Stufe 1 angab											
Kosten in Euro für Teiln. von Typ A											
Teiln. von Typ B ist Raucher											

Mit Hilfe der von Ihnen in diesem Fragebogen beschriebenen Strategie wird ein weiteres Experiment durchgeführt.

Appendix B

Do people have a preference for increasing or decreasing pain? An experimental comparison of hypothetical and monetary consequences

Written experimental instructions

Anleitung zum Experiment

Vielen Dank für Ihre Teilnahme an unserem Experiment.

Sie nehmen heute an einem Experiment teil, in dem es um mehrere Entscheidungen geht. In diesem Experiment gibt es keine richtigen oder falschen Antworten.

Zu Beginn des Experiments werden Sie zwei Schmerz-Erfahrungen machen. Hierfür müssen Sie Ihre Hand für einige Zeit in kaltes Wasser eintauchen. Jede Erfahrung besteht dabei aus drei verschiedenen kalten Wasserbecken.

Anschließend treffen Sie verschiedene Entscheidungen, die sich auf eine Wiederholung der Schmerz-Erfahrung mit den Kaltwasserbecken beziehen. Hierfür bekommen Sie zwei Antwortbögen, auf dem ersten treffen Sie eine Entscheidung, auf dem folgenden die restlichen.

Am Ende wird eine Ihrer Entscheidungen in die Tat umgesetzt. Um zu bestimmen, welche Entscheidung das ist, ziehen Sie zufällig ein Los aus einer Urne. Die Urne enthält ein Los für jede Ihrer Entscheidungen. Die am Ende mit den Losen ermittelte Entscheidung steht für eine Schmerzerfahrung, die dann umgesetzt wird.

Ihre erste Entscheidung:

Welche der beiden eben gemachten Schmerzerfahrungen möchten Sie lieber wiederholen? Bitte setzen Sie ein Kreuz für eine der beiden Erfahrungen oder bei „egal“, für den Fall, dass Sie indifferent zwischen beiden Erfahrungen sind.

			A	B	egal
1	A: Wiederholung der ersten Erfahrung	B: Wiederholung der zweiten Erfahrung.			

Als nächstes geht es um die spätere Wiederholung einer der beiden Erfahrungen, die Sie eben gemacht haben. Sie werden nun einen Entscheidungsbogen mit mehreren Entscheidungen bekommen. Nachfolgend haben Sie immer die Wahl zwischen drei Alternativen:

1 Sie möchten lieber Alternative A (die eine von den beiden Erfahrungen).

2 Sie möchten lieber Alternative B (die andere von den beiden Erfahrung und zusätzlich einen Geldbetrag an uns bezahlen).

3 Es ist Ihnen egal, ob Sie Alternative A oder Alternative B bekommen.

Bitte kreuzen Sie nachfolgend immer an, was Sie wählen.

Nr.	**Alternative A**	**Alternative B**	**A**	**B**	**egal**
2	2. Erfahrung	1. Erfahrung und Zahlung von 0,20 Euro			
3	2. Erfahrung	1. Erfahrung und Zahlung von 0,40 Euro			
4	2. Erfahrung	1. Erfahrung und Zahlung von 0,60 Euro			
5	2. Erfahrung	1. Erfahrung und Zahlung von 0,80 Euro			
6	2. Erfahrung	1. Erfahrung und Zahlung von 1,00 Euro			
7	2. Erfahrung	1. Erfahrung und Zahlung von 1,20 Euro			
8	2. Erfahrung	1. Erfahrung und Zahlung von 1,40 Euro			
9	2. Erfahrung	1. Erfahrung und Zahlung von 1,60 Euro			
10	2. Erfahrung	1. Erfahrung und Zahlung von 1,80 Euro			
11	2. Erfahrung	1. Erfahrung und Zahlung von 2,00 Euro			
12	2. Erfahrung	1. Erfahrung und Zahlung von 2,20 Euro			
13	2. Erfahrung	1. Erfahrung und Zahlung von 2,40 Euro			
14	2. Erfahrung	1. Erfahrung und Zahlung von 2,60 Euro			
15	2. Erfahrung	1. Erfahrung und Zahlung von 2,80 Euro			
16	2. Erfahrung	1. Erfahrung und Zahlung von 3,00 Euro			
17	2. Erfahrung	1. Erfahrung und Zahlung von 3,20 Euro			
18	2. Erfahrung	1. Erfahrung und Zahlung von 3,40 Euro			
19	2. Erfahrung	1. Erfahrung und Zahlung von 3,60 Euro			
20	2. Erfahrung	1. Erfahrung und Zahlung von 3,80 Euro			
21	2. Erfahrung	1. Erfahrung und Zahlung von 4,00 Euro			
22	2. Erfahrung	1. Erfahrung und Zahlung von 4,20 Euro			
23	2. Erfahrung	1. Erfahrung und Zahlung von 4,40 Euro			
24	2. Erfahrung	1. Erfahrung und Zahlung von 4,60 Euro			
25	2. Erfahrung	1. Erfahrung und Zahlung von 4,80 Euro			
26	2. Erfahrung	1. Erfahrung und Zahlung von 5,00 Euro			

Nr.	Alternative A	Alternative B	A	B	egal
27	2. Erfahrung und Zahlung von 0,20 Euro	1. Erfahrung			
28	2. Erfahrung und Zahlung von 0,40 Euro	1. Erfahrung			
29	2. Erfahrung und Zahlung von 0,60 Euro	1. Erfahrung			
30	2. Erfahrung und Zahlung von 0,80 Euro	1. Erfahrung			
31	2. Erfahrung und Zahlung von 1,00 Euro	1. Erfahrung			
32	2. Erfahrung und Zahlung von 1,20 Euro	1. Erfahrung			
33	2. Erfahrung und Zahlung von 1,40 Euro	1. Erfahrung			
34	2. Erfahrung und Zahlung von 1,60 Euro	1. Erfahrung			
35	2. Erfahrung und Zahlung von 1,80 Euro	1. Erfahrung			
36	2. Erfahrung und Zahlung von 2,00 Euro	1. Erfahrung			
37	2. Erfahrung und Zahlung von 2,20 Euro	1. Erfahrung			
38	2. Erfahrung und Zahlung von 2,40 Euro	1. Erfahrung			
39	2. Erfahrung und Zahlung von 2,60 Euro	1. Erfahrung			
40	2. Erfahrung und Zahlung von 2,80 Euro	1. Erfahrung			
41	2. Erfahrung und Zahlung von 3,00 Euro	1. Erfahrung			
42	2. Erfahrung und Zahlung von 3,20 Euro	1. Erfahrung			
43	2. Erfahrung und Zahlung von 3,40 Euro	1. Erfahrung			
44	2. Erfahrung und Zahlung von 3,60 Euro	1. Erfahrung			
45	2. Erfahrung und Zahlung von 3,80 Euro	1. Erfahrung			
46	2. Erfahrung und Zahlung von 4,00 Euro	1. Erfahrung			
47	2. Erfahrung und Zahlung von 4,20 Euro	1. Erfahrung			
48	2. Erfahrung und Zahlung von 4,40 Euro	1. Erfahrung			
49	2. Erfahrung und Zahlung von 4,60 Euro	1. Erfahrung			
50	2. Erfahrung und Zahlung von 4,80 Euro	1. Erfahrung			
51	2. Erfahrung und Zahlung von 5,00 Euro	1. Erfahrung			

Nr.	Alternative A	Alternative B	A	B	egal
2	1. Erfahrung	2. Erfahrung und Zahlung von 0,20 Euro			
3	1. Erfahrung	2. Erfahrung und Zahlung von 0,40 Euro			
4	1. Erfahrung	2. Erfahrung und Zahlung von 0,60 Euro			
5	1. Erfahrung	2. Erfahrung und Zahlung von 0,80 Euro			
6	1. Erfahrung	2. Erfahrung und Zahlung von 1,00 Euro			
7	1. Erfahrung	2. Erfahrung und Zahlung von 1,20 Euro			
8	1. Erfahrung	2. Erfahrung und Zahlung von 1,40 Euro			
9	1. Erfahrung	2. Erfahrung und Zahlung von 1,60 Euro			
10	1. Erfahrung	2. Erfahrung und Zahlung von 1,80 Euro			
11	1. Erfahrung	2. Erfahrung und Zahlung von 2,00 Euro			
12	1. Erfahrung	2. Erfahrung und Zahlung von 2,20 Euro			
13	1. Erfahrung	2. Erfahrung und Zahlung von 2,40 Euro			
14	1. Erfahrung	2. Erfahrung und Zahlung von 2,60 Euro			
15	1. Erfahrung	2. Erfahrung und Zahlung von 2,80 Euro			
16	1. Erfahrung	2. Erfahrung und Zahlung von 3,00 Euro			
17	1. Erfahrung	2. Erfahrung und Zahlung von 3,20 Euro			
18	1. Erfahrung	2. Erfahrung und Zahlung von 3,40 Euro			
19	1. Erfahrung	2. Erfahrung und Zahlung von 3,60 Euro			
20	1. Erfahrung	2. Erfahrung und Zahlung von 3,80 Euro			
21	1. Erfahrung	2. Erfahrung und Zahlung von 4,00 Euro			
22	1. Erfahrung	2. Erfahrung und Zahlung von 4,20 Euro			
23	1. Erfahrung	2. Erfahrung und Zahlung von 4,40 Euro			
24	1. Erfahrung	2. Erfahrung und Zahlung von 4,60 Euro			
25	1. Erfahrung	2. Erfahrung und Zahlung von 4,80 Euro			
26	1. Erfahrung	2. Erfahrung und Zahlung von 5,00 Euro			

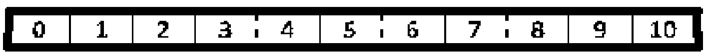
Nr.	Alternative A	Alternative B	A	B	egal
27	1. Erfahrung und Zahlung von 0,20 Euro	2. Erfahrung			
28	1. Erfahrung und Zahlung von 0,40 Euro	2. Erfahrung			
29	1. Erfahrung und Zahlung von 0,60 Euro	2. Erfahrung			
30	1. Erfahrung und Zahlung von 0,80 Euro	2. Erfahrung			
31	1. Erfahrung und Zahlung von 1,00 Euro	2. Erfahrung			
32	1. Erfahrung und Zahlung von 1,20 Euro	2. Erfahrung			
33	1. Erfahrung und Zahlung von 1,40 Euro	2. Erfahrung			
34	1. Erfahrung und Zahlung von 1,60 Euro	2. Erfahrung			
35	1. Erfahrung und Zahlung von 1,80 Euro	2. Erfahrung			
36	1. Erfahrung und Zahlung von 2,00 Euro	2. Erfahrung			
37	1. Erfahrung und Zahlung von 2,20 Euro	2. Erfahrung			
38	1. Erfahrung und Zahlung von 2,40 Euro	2. Erfahrung			
39	1. Erfahrung und Zahlung von 2,60 Euro	2. Erfahrung			
40	1. Erfahrung und Zahlung von 2,80 Euro	2. Erfahrung			
41	1. Erfahrung und Zahlung von 3,00 Euro	2. Erfahrung			
42	1. Erfahrung und Zahlung von 3,20 Euro	2. Erfahrung			
43	1. Erfahrung und Zahlung von 3,40 Euro	2. Erfahrung			
44	1. Erfahrung und Zahlung von 3,60 Euro	2. Erfahrung			
45	1. Erfahrung und Zahlung von 3,80 Euro	2. Erfahrung			
46	1. Erfahrung und Zahlung von 4,00 Euro	2. Erfahrung			
47	1. Erfahrung und Zahlung von 4,20 Euro	2. Erfahrung			
48	1. Erfahrung und Zahlung von 4,40 Euro	2. Erfahrung			
49	1. Erfahrung und Zahlung von 4,60 Euro	2. Erfahrung			
50	1. Erfahrung und Zahlung von 4,80 Euro	2. Erfahrung			
51	1. Erfahrung und Zahlung von 5,00 Euro	2. Erfahrung			

Bitte bewerten Sie die beiden Schmerzerfahrungen 1 und 2 jeweils getrennt auf den beiden folgenden Skalen. Kreuzen Sie dabei die Zahl auf dem ersten Balken an, die Ihrer erlebten Schmerzintensität bei der ersten Erfahrung entspricht. Die Null steht in diesem Fall für keinen Schmerz, während die Zehn für den stärksten vorstellbaren Schmerz steht.

Bitte kreuzen Sie genauso auch auf der zweiten Skala die Zahl an, die Ihrer wahrgenommenen Schmerzintensität bei der zweiten Erfahrung entspricht.

Erfahrung 1

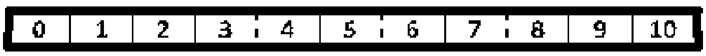
Kein Schmerz



Stärkster vorstellbarer Schmerz

Erfahrung 2

Kein Schmerz



Stärkster vorstellbarer Schmerz

Appendix C

Determining risk preferences for pain

Written experimental instructions

Anleitung zum Experiment

Vielen Dank für Ihre Teilnahme an unserem Experiment.

Sie nehmen heute an einem Experiment teil, in dem es um zehn Lotterieentscheidungen geht. In diesem Experiment gibt es keine richtigen oder falschen Antworten. Vor den Entscheidungen wird an einem Beispiel erklärt, wie genau Sie Ihre Antworten geben können.

Am Ende wird per Zufall bestimmt, welche Ihrer Entscheidungen tatsächlich ausgespielt wird. Das heißt, durch Ziehen einer von zehn Kugeln aus einer Urne wird dies festgelegt. Die so ausgeloste Lotterie wird dann mit Kugeln zweier Farben, verteilt nach der zugehörigen Wahrscheinlichkeit, ausgespielt.

Ablauf des Experiments

Die Entscheidungen werden zwischen Lotterien getroffen, die eine Kaltwasser-Aufgabe beinhalten. Dabei wird eine Hand bis zum Handgelenk in kaltes Wasser eingetaucht. Zentral sind dabei vier Becken mit unterschiedlich kaltem Wasser: 14°, 10°, 7° und 4° Celsius.

Die zehn Lotterieentscheidungen beinhalten die Wassertemperaturen der Becken. Die am Ende ermittelte Lotterie entscheidet, in welches Becken eine Hand für 2 Minuten eingetaucht werden muss.

Um die Temperatur der Becken vor dem Treffen der Entscheidungen einschätzen zu können, werden alle vier Becken vor dem Experiment ausprobiert. Der Ablauf ist wie folgt:

- Die eine Hand wird für zwei Minuten in das erste Becken (14°) eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die andere Hand wird für zwei Minuten in das nächste Becken (10°) eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die erste Hand wird für zwei Minuten in das dritte Becken (7°) eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die andere Hand wird für zwei Minuten in das vierte Becken (4°) eingetaucht.

Die Experimentalleiterin gibt die Zeiten an.

Nach dem Ausprobieren werden die zehn Entscheidungen getroffen.

Die Lotterien

Inhalt dieses Experiments sind Lotterien. Sie spielen in diesem Fall aber nicht um einen Gewinn, sondern in jeder Lotterie um einen von zwei verschiedenen Ausgängen. Diese Ausgänge bestehen in einer Kaltwasseraufgabe. Dies bedeutet, Sie müssen eine Hand für zwei Minuten in kaltes Wasser eintauchen. Welche Temperatur dieses Wasser hat, wird durch die Entscheidungen in der Lotterie ermittelt.

Als erstes bitten wir Sie, aus zwei Lotterien zu wählen. Diese Lotterien sind dabei alle von folgendem Typ:

Wahrscheinlichkeit	p%	(100-p)%
Wassertemperatur beim Eintauchen für zwei Minuten	Temperatur 1	Temperatur 2

Dabei bezeichnet p die Wahrscheinlichkeit, mit der Sie die Hand für zwei Minuten bei Temperatur 1 eintauchen und (100-p) die Wahrscheinlichkeit, mit der Sie die Hand bei Temperatur 2 zwei Minuten lang eintauchen.

Im Beispiel

Wahrscheinlichkeit	90%	10%
Wassertemperatur beim Eintauchen für zwei Minuten (in Grad Celsius)	14°	4°

müssen Sie mit der Wahrscheinlichkeit von 90% Ihre Hand für zwei Minuten in 14° Celsius kaltes Wasser halten und mit der Wahrscheinlichkeit von 10% für zwei Minuten in 4° Celsius kaltes Wasser. Nur eine der beiden Optionen wird umgesetzt.

Die Auswahl der Temperaturen erfolgt durch Ziehen einer Kugel aus einer Urne mit 10 Kugeln. Dabei sind p (im Beispiel 9) rot und 100-p (im Beispiel 1) blau. Eine Kugel wird gezogen. Bei rot müssen Sie Ihre Hand für zwei Minuten in 14° kaltes Wasser halten, bei blau für zwei Minuten in 4° kaltes Wasser.

Bitte wählen Sie im Folgenden immer zwischen den zwei angebotenen Alternativen:

A für Alternative A,

B für Alternative B.

1		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	10%	90%	10%	90%		
	Eintauchen für 2 min	14°	4°	10°	7°		

2		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	20%	80%	20%	80%		
	Eintauchen für 2 min	14°	4°	10°	7°		

3		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	30%	70%	30%	70%		
	Eintauchen für 2 min	14°	4°	10°	7°		

4		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	40%	60%	40%	60%		
	Eintauchen für 2 min	14°	4°	10°	7°		

5		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	50%	50%	50%	50%		
	Eintauchen für 2 min	14°	4°	10°	7°		

6		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	60%	40%	60%	40%		
	Eintauchen für 2 min	14°	4°	10°	7°		

7		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	70%	30%	70%	30%		
	Eintauchen für 2 min	14°	4°	10°	7°		

8		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	80%	20%	80%	20%		
	Eintauchen für 2 min	14°	4°	10°	7°		

9		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	90%	10%	90%	10%		
	Eintauchen für 2 min	14°	4°	10°	7°		

10		Alternative A		Alternative B		A	B
	Wahrscheinlichkeit	100%	0%	100%	0%		
	Eintauchen für 2 min	14°	4°	10°	7°		

Ablauf des Experiments

Die Entscheidungen werden zwischen Lotterien getroffen, die eine Kaltwasser-Aufgabe beinhalten. Dabei wird eine Hand bis zum Handgelenk in kaltes Wasser eingetaucht. Zentral sind dabei vier Eintauchdauern in ein Becken mit 4° Celsius kaltem Wasser: 2, 4, 8 und 12 Minuten.

Um die Temperatur des Beckens von 4° vor dem Treffen der Entscheidungen einschätzen zu können, wird dieses vor dem Experiment für zwei Minuten ausprobiert.

Die zehn Lotterieentscheidungen beinhalten die Eintauchdauern in das 4° kalte Becken. Die am Ende ermittelte Lotterie entscheidet, wie lange Ihre Hand in das Becken eingetaucht werden muss. Dabei bleibt die Hand nie länger als zwei Minuten am Stück eingetaucht.

Falls zum Beispiel eine Eintauchdauer von 8 Minuten in der Lotterie ermittelt wird, so liefe die Realisierung am Ende des Experiments wie folgt ab:

- Die rechte Hand wird für zwei Minuten in das Becken eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die linke Hand wird für zwei Minuten in das Becken eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die rechte Hand wird für zwei Minuten in das Becken eingetaucht.
  Danach gibt es eine Pause von 30 Sekunden.
- Die linke Hand wird für zwei Minuten in das Becken eingetaucht.

Nach dem Ausprobieren des Beckens durch 2-minütiges Eintauchen werden die zehn Entscheidungen getroffen.

Die Lotterien

Inhalt dieses Experiments sind Lotterien. Sie spielen in diesem Fall aber nicht um einen Gewinn, sondern in jeder Lotterie um einen von zwei verschiedenen Ausgängen. Diese Ausgänge bestehen in einer Kaltwasseraufgabe. Dies bedeutet, Sie müssen eine Hand für eine bestimmte Zeit in 4° Celsius kaltes Wasser halten. Wie lange Sie die Hand eingetaucht lassen müssen, wird durch die Ergebnisse in der Lotterie ermittelt.

Als erstes bitten wir Sie, aus zwei Lotterien zu wählen, diese Lotterien sind dabei alle von folgendem Typ:

Wahrscheinlichkeit	P%	(100-p)%
Eintauchzeit bei 4°	T1	T2

Dabei bezeichnet p die Wahrscheinlichkeit, mit der Sie die Hand T1 Minuten eintauchen und (100-p) die Wahrscheinlichkeit, mit der Sie die Hand T2 Minuten lang eintauchen.

Im Beispiel

Wahrscheinlichkeit	90%	10%
Eintauchzeit bei 4° Celsius	2 x 2 Minuten	4 x 2 Minuten

müssen Sie mit der Wahrscheinlichkeit von 90% Ihre Hand für 2 mal 2 Minuten in 4° kaltes Wasser halten und mit der Wahrscheinlichkeit von 10% für 4 mal 2 Minuten. Nur eine der beiden Optionen wird umgesetzt. Eine Eintauchzeit ist dabei nie länger als 2 Minuten, jedoch variiert die Anzahl der Eintauchdurchgänge in den Lotterien, zum Beispiel 4 mal 2 Minuten.

Die Auswahl der Eintauchdauer erfolgt durch Ziehen einer Kugel aus einer Urne mit 10 Kugeln. Dabei sind p (im Beispiel 9) rot und 100-p (im Beispiel 1) blau. Eine Kugel wird gezogen. Bei rot müssen Sie Ihre Hand 2 mal 2 Minuten in 4° kaltes Wasser halten, bei blau für 4 mal 2 Minuten.

Bitte wählen Sie im Folgenden immer zwischen den zwei angebotenen Alternativen:

A für Alternative A,

B für Alternative B.

Bitte geben Sie auf dem Lotteriebogen im oberen Kasten Ihr Geschlecht an.

		Alternative A		Alternative B		A	B
1	Wahrscheinlichkeit	10%	90%	10%	90%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
2	Wahrscheinlichkeit	20%	80%	20%	80%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
3	Wahrscheinlichkeit	30%	70%	30%	70%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
4	Wahrscheinlichkeit	40%	60%	40%	60%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
5	Wahrscheinlichkeit	50%	50%	50%	50%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
6	Wahrscheinlichkeit	60%	40%	60%	40%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
7	Wahrscheinlichkeit	70%	30%	70%	30%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
8	Wahrscheinlichkeit	80%	20%	80%	20%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
9	Wahrscheinlichkeit	90%	10%	90%	10%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		

		Alternative A		Alternative B		A	B
10	Wahrscheinlichkeit	100%	0%	100%	0%		
	Eintauchzeit bei 4°	1 x 2 min	6 x 2 min	2 x 2 min	4 x 2 min		