

Thomas Schelling und die Theorie der Self-Command

Andreas Ortmann¹ & Angelika Weber²

Erscheint in:

Ingo Pies & Martin Leschke, Thomas Schellings strategische Ökonomik,
Mohr-Siebeck, Tübingen

¹ CERGE-EI, Karls-Universität und Akademie der Wissenschaften der Tschechischen Republik, Prag

² Fachhochschule Würzburg-Schweinfurt

1. Einführung

Was ist unter Self-Command zu verstehen? Und warum ist es überhaupt ein Thema für die Ökonomie — ein Thema mit dem sich selbst so hochrangige Wissenschaftler wie Schelling befassen? Was ist der Kern von Schelling's Theorie der Self-Command? Self-Command wird auch als Fähigkeit zur Selbstkontrolle, Selbstdisziplin, Selbstbeherrschung, Selbstbindung, Selbstregulation oder als Willenskraft beschrieben¹. Um Ziele zu erreichen, müssen wir unser Verhalten (bewusst oder unbewusst) zielführend steuern. Schwierig wird es immer dann, wenn zum Zeitpunkt der Ausführung eines Verhaltens ein Bedürfnis im Vordergrund steht, das dem zielführenden Verhalten entgegensteht und wir uns überwinden müssen, dennoch das zu tun, was wir vorher beschlossen haben. Dieser Konflikt zieht sich durch alle Bereiche des Alltags, betrifft aber auch ethische, rechtliche und gesellschaftspolitische Fragen mit entsprechenden Konsequenzen. Thomas Schelling hat sich intensiv und in kreativer Weise mit Problemen von Self-Command auseinandergesetzt, also mit jenen Konfliktsituationen die uns eine Beherrschung unserer aktuellen Bedürfnisse abverlangen (und an der wir oft scheitern).

Der folgende Beitrag befasst sich zunächst damit, was Thomas Schelling über Probleme von Self-Command zu sagen hat. Nach einer Zusammenfassung seiner Theorie der Self-Command wird diese in einem größeren Kontext betrachtet und kritisch bewertet.

2. Schelling's Theorie der Self-Command

Warum ist es für viele von uns so schwierig, sich hinzusetzen und ein Buchkapitel wie dieses zu schreiben, obwohl wir uns längst und durchaus bewusst entschieden haben, es zu schreiben? Oder: Warum haben viele von uns Schwierigkeiten, regelmäßig Sport zu treiben (oder irgendeine andere Aktivität, die uns wahrscheinlich auf lange Sicht Nutzen bringt), obwohl wir uns eigentlich dafür entschieden haben es zu tun? Und: Warum haben so viele Schwierigkeiten, das Rauchen aufzugeben (oder irgendein anderes Verhalten, das uns womöglich langfristig teuer zu stehen kommt) obwohl sie glaubhaft versichern, es ernsthaft zu wollen?

¹ Um Missverständnisse zu vermeiden, verwenden wir im Text den englischen Ausdruck.

Schelling analysierte diese und zahlreiche ähnliche Probleme von Self-Command in einer Reihe von Beiträgen. Die in dieser Hinsicht bedeutendste und leicht zugängliche Veröffentlichung ist Schelling (1984). Sie stützt sich stark auf frühere Arbeiten von Schelling (1980, 1982), die in Schelling (1984a) nachgedruckt wurden. Bedeutsam ist, dass Schelling dabei auch über Taktiken und Techniken des „Selbstmanagement“ (Schelling, 1980; S.63) – oder “strategic egonomics” wie er es benennt (Schelling, 1982; S. 63) – nachgedacht hat. Es sind Taktiken und Techniken, die wir benutzen, um uns selbst zu überlisten (z.B. um dieses Kapitel zu schreiben, Sport zu treiben, das Rauchen aufzugeben...)².

Uns selbst überlisten? „Selbstmanagement“? „Strategic egonomics“? Schelling nimmt an, dass sich Menschen manchmal so verhalten, als ob sie zwei Selbst hätten, die miteinander im Konflikt stehen: “People behave sometimes as if they have two selves, one who wants clean lungs and long life and another who adores tobacco, or one who wants a lean body and another who wants dessert, or one who yearns to improve himself by reading Adam Smith on self-command (in *Theory of Moral Sentiments*) and another who would rather watch an old movie on television. The two are in continual contest for control.” (Schelling, 1980; S. 58).

In dem er das Problem von Self-Command als ein Problem eines ständigen Kampfs um Kontrolle zwischen (mindestens) zwei Selbst konzipiert, hat sich Schelling – so wie einige andere vor ihm – von der “individualistic-utilitarian foundation of neo-classical economics” (Schelling, 1980; S. 59), die Menschen³ als Einheiten sieht, entfernt: “If we accept the idea of two selves of which usually one is in charge at a time, or two value systems that are alternate rather than subject to simultaneous and integrated scrutiny, ‘rational decision’ has to be replaced with something like collective choice. Two or more selves that alternatively occupy the same individual that have different goals and tastes, ... have to be construed as engaged not in joint optimization but in a strategic game.”

² Schelling möchte seine Ideen dabei zwar von den Begriffen Selbst-Kontrolle oder Selbstdisziplin, wie sie üblicherweise verstanden werden, abgrenzen (Schelling, 1980): „What I’m talking about is different from what is usually thought of as self-control or self-discipline. I am not talking about the development of inner strength, character, or moral fiber, or the change in values that goes with religious conversion“ (S. 69), jedoch gibt er selbst keine klare Definition von Self-Command.

³ Aber auch Familien und sogar Firmen und Organisationen, wie Schelling anmerkte (1980, S. 60).

(Schelling, 1982; S. 93 f.). Wenn also zwei (oder mehr) Selbst sozusagen um die Vorherrschaft ringen, ist, als Konsequenz, dieser Prozess in den Augen von Schelling als strategisches Spiel aufzufassen.

In diesem strategischen Spiel entsteht ein Problem der Self-Command immer dann, wenn Präferenzen und Werte der zwei Selbst inkompatibel sind und wenn ein Verhalten – wie z.B. Sport zu treiben, oder keine Zigarette anzuzünden, oder den ersten Drink abzulehnen – freiwillig und bewusst (“voluntary and conscious”, Schelling, 1980; S. 64) ist, aber nicht ausgeführt werden kann.

Es ist diese Konzeption von zwei, in der Sprache der modernen Ökonomie *zeit-inkonsistenten* Selbst, die den Kern von Schelling’s Theorie der Self-Command bildet. Vor dem Hintergrund dieser Sichtweise werden Fragen aufgeworfen, die notorisch schwierig zu beantworten sind: Wie bewerten wir (unsere) Präferenzen heute und wie bewerten wir sie morgen? Mit Schelling’s extremsten Beispiel (1984; S. 2)⁴ illustriert: Wie bewerten wir den Wunsch einer Person heute vor Selbstmord geschützt zu werden, wenn die gleiche Person unter anderen (aber vorhersehbaren) Umständen morgen klar bevorzugt, tot zu sein? Oder wie bewerten wir den Wunsch von jemanden heute zu sterben, wenn diese Person nicht in der Lage ist „...to face the finality of bringing it [death] about, and, like the parachutist who asks to be shoved out if he grips the door jam, implores our help in getting him over the brink” (Schelling, 1984; S. 2)? Oder, um ein weniger dramatisches, aber immer wiederkehrendes Beispiel in Schelling’s Reflektionen über Self-Command (z.B. Schelling, 1980, S. 73 ff.; oder Schelling, 1992) zu zitieren: Wie ist der unmittelbare Genuss durch das Rauchen gegenüber dem möglichen, zukünftigen gesundheitlichen Gewinn durch die Beendigung dieser Sucht zu bewerten?

In jedem seiner Artikel über Self-Command, sucht Schelling Antworten zu diesen und ähnlichen Fragen von Wohlfahrt. Gelegentlich (siehe das Beispiel Captain Ahab⁵) äußert

⁴ “There is no graver issue for the coming century than how to recognize and authenticate the preferences of people for whom dying has become the issue that dominates their lives. This is the ultimate dilemma of authenticating the self, of discovering the legitimate sovereignty of the individual.” (Schelling, 1982; S. 112)

⁵ “A few years ago I saw again the original Moby Dick There was a scene ... of Ahab in the water losing his leg, and immediately afterward below deck under a blanket, eating an apple with three of the crew. The blacksmith enters with a hot iron to cauterize the stump. ... When I was first contemplating this

er dabei aber auch Zweifel über seine früheren Einschätzungen über Nutzen und Kosten bestimmter Aktionen. Was ist das „wahre“ (Schelling, 1984; S. 2) oder das „authentische“ (Schelling, 1980; S. 5) Selbst? Gibt es ein solches Selbst?⁶ Existiert das authentische Selbst per se?⁷

Abgesehen von den neuen und schwierigen Überlegungen zur Identität und zur Wohlfahrt, die durch das Problem von sich widersprechenden Präferenzen und Werten aufgeworfen werden, werden durch die Analyse der Self-Command zusätzlich interessante rechtliche, ethische und politische Fragen angeregt: “Legal issues arise in some attempts to abdicate rights that are deemed to be inalienable. I cannot get a court injunction against my own smoking. I cannot contract with a skydiving pilot to push me out of the airplane. ... I cannot contract with a fat farm to hold me against my will until I have lost some number of pounds; they have to let me out when I ask.” (Schelling, 1984; S. 2). Sollte es solche gerichtlichen Verfügungen geben? Sollten solche Möglichkeiten eines Vertrags existieren? Ist es richtig und ethisch vertretbar, dass das Recht sich auf die Seite des Selbst stellt, welches nicht sterben will? Sollte das Gesetz Partei ergreifen, wenn das authentische Selbst nicht in Frage steht?⁸

Wie Zeckhauser (1988) richtig bemerkt, stellt sich immer das folgende Problem: “...if I can not choose for myself, then on what basis can others decide that I should not be afforded this freedom?” (S. 161). Und diese Frage wird umso mehr relevant und drängend, wenn, wie im Fall von Rauchen, eine weitere Partei vom Konflikt des (authentischen) Selbst mit dem anderen Selbst betroffen ist.

Während Schelling aufgrund der Konzeption unseres Selbst als zwei zeit-inkonsistente

episode I thought it an incontestable case of the utility gain from denying freedom of choice and ignoring revealed preference. ... how do we know whether an hour of extreme pain is more than life is worth? Alternatively, how do we know whether an hour of extreme pain is more than death is worth? The conclusion that I reach is that I do not know, not for you and not for me.” (Schelling, 1984; S. 9)

⁶ “The question, which is the authentic one, may define the problem wrong. Both selves can be authentic.” (Schelling, 1984; S. 9; siehe auch fast wörtlich Schelling, 1982; S. 108)

⁷ Was eher unwahrscheinlich ist (siehe LeDoux, 2002).

⁸ “Anyone who is happily addicted to nicotine, benzedrine, valium, chocolate, heroin, or horse racing, and anyone unhappily addicted who would not elect the pains and deprivations of withdrawal, are not my subject. I’m not concerned with whether cigarettes or rich desserts are bad for you, only with the fact that there are people who wish so badly to avoid them that, if they could, they would put these commodities beyond their own reach.” (Schelling, 1984; S. 4 f.)

Selbst mit Fragen der Wohlfahrt sowie ethischen, rechtlichen und politischen Fragen rang, war das Hauptanliegen seiner Diskussion stets, brillante Lösungen zu finden, die keine rechtlichen Instrumente bzw. ethische oder politische Diskussionen notwendig machen. Mit anderen Worten, er war daran interessiert, *zeit-konsistente* Mechanismen oder Selbstbindungsstrategien zu finden, die es, entsprechend ausgeführt, dem heutigen Selbst erlauben, die Auswahlmöglichkeiten eines zukünftigen Selbst in einer Weise einzuschränken, dass es im Ergebnis wie das heutige Selbst handeln wird: “If we are clever we can arrange it; I go to a remote fat farm that requires a 24-hour notice to order a car, a notice that I can rescind during a moment’s resurgent resolve to lose weight.” (Schelling, 1984; S. 2). Leider sind solche Lösungen nicht immer einfach zu finden: Wie ist es möglich, sich von Zigaretten fern zu halten, wenn man sie an jeder Ecke kaufen kann? Sollte ich mich ohne Zigaretten in der Wildnis absetzen lassen, so dass es Tage dauert, bis ich wieder in die Nähe von Zigarettenautomaten komme?

Um Lösungen zu finden, schlägt Schelling vor, uns auf Prinzipien des Umgangs miteinander (insbesondere mit den Personen, die uns nahe stehen) zu besinnen und zu beherzigen: “Many of the skills and maxims and stratagems for coping with one’s own behavior become less mystifying and more familiar if we can recognize them as the same principles and stratagems that apply to managing someone else – someone in a close relation, with a paternalist or senior-junior quality like that between parent and child, teacher and pupil, missionary and convert, master and apprentice, or guide and follower.” (Schelling, 1980; S. 63). Schelling nimmt also an, dass uns Lernen am Modell beim Finden von effektiven Selbstbindungsstrategien unterstützt.

Seine Veröffentlichungen (z.B. Schelling 1984; S. 6 f.) enthalten zahlreiche Vorschläge für Self-Command-Strategien, die es dem authentischen Selbst ermöglichen sollten, die Oberhand zu behalten. Einige seiner Beispiele sind erprobte und bewährte Methoden, die dem Leser dieses Kapitels durchaus vertraut sein dürften: “Relinquish authority to somebody else: let him hold your car keys. ... Remove the mischievous resources: don’t keep liquor, or sleeping pills, in the house; order a hotel room without television. ... Incarcerate yourself: Have somebody drop you at a cheap motel without telephone or television and call for you after eight hours’ work. ... Arrange delays: the crisis may pass before the time is up. Use buddies and teams: exercise together, order each other’s

lunches. Automate the behavior. The automation that I look forward to is a device implanted to monitor cerebral hemorrhage that, if the stroke is severe enough to indicate a hideous survival, kills the patient before anyone can intervene to remove it.” (S. 6 f.)

3. Kritische Würdigung von Schelling’s Theorie der Self-Command

Thomas Schelling’s Theorie der Self-Command (z.B. Schelling; 1984a) hatte weniger Einfluss als seine Überlegungen zur Strategie des Konflikts (Schelling; 1980)⁹, obwohl die „Strategies of commitment“ – so der sehr treffende Titel einer kürzlich erschienen Sammlung seiner Essays (Schelling; 2006) – für (vielleicht) nicht zeit-konsistente Aktionen, ein gemeinsames Thema beider Bücher ist.

Wir haben drei Vermutungen, warum dem so ist. Ein Grund für den geringeren Einfluss von Schelling’s Arbeiten über Self-Command mag sein, dass diese Beiträge eher in einer Arat Plauderton geschrieben sind und – obwohl oft wunderbar unterhaltsam und kreativ in den Beispielen und Vorschlägen für Bewältigungsstrategien und Taktiken zur Selbstbindung – den wissenschaftlich interessierten Leser nicht mit einem theoretischem Rahmen versorgt, der es erlaubt, Zusammenhänge zwischen einzelnen Erkenntnissen herzustellen und theoretisch einzuordnen. Thomas Schelling, ähnlich wie Adam Smith, hat die Fähigkeit, dem Leser interessante Beispiele zu liefern, auf die er sich mit seinen theoretischen Überlegungen beziehen kann. Schelling machte zwar wiederholt Versuche eine Klassifikationen der Art und Umstände, in denen Self-Command versagen könnte zu erstellen, jedoch führten diese Versuche zu keiner brauchbaren Taxonomie. Obwohl streckenweise durchaus unterhaltsam, kann der erzählende, nicht-mathematische Stil für den Schelling berühmt (berüchtigt) ist, auch zu frustrierendem Lesen und frustrierten Lesern führen.

Schelling ist sich dieser Schwäche durchaus bewusst: “It is one thing to appreciate the general idea of tactics deployed to protect oneself from oneself, and the ubiquitousness of

⁹ Eine Suchanfrage auf „scholar.google“ Anfang July 2006 quantifiziert diese Aussage: Während Schelling’s *The Strategy of Conflict* (1960) und *Micromotives and Macrobehavior* (1978) fast 2000 bzw. mehr als 1000 Einträge aufweisen, finden sich für *Choice and Consequence* (1984) lediglich 150 Zitationen. Selbst wenn man den Zeitfaktor kontrolliert, steht außer Frage, dass seine Arbeiten über Self-Command und verwandte Themen weniger Eindruck auf die Profession gemacht haben.

the problem. It is something quite different to focus on a specific problem, and to do so not to illuminate a general principle but to cope with the mischief at hand. For that we need a systematic way of analyzing the habit or weakness along relevant dimensions: the vulnerabilities of its victim, the environment of its victim, the environment in which it occurs, and the information, communication, and institutional commitments that be brought to bear. I do not know of any taxonomy or analytical scheme for finding the similarities and highlighting the differences among the different habits or addictions and the targets they afflict. I can only illustrate the kinds of analytical dimensions I have in mind.” (Schelling, 1980; S. 69 f.).

Ein zweiter Grund für den geringeren Einfluss von Schelling’s Arbeit über Self-Command könnte sein, dass einige der Ideen nicht so sind, wie es Schelling darzustellen versuchte. Insbesondere Strotz (1956; ebenso wie Ainslee, 1975) hatte fast zwanzig Jahre zuvor die Frage von Zeitinkonsistenz in der Ökonomie aufgegriffen (seitdem wurde sein Artikel eine symbolische Referenz für jeden, der über dieses Thema schreibt). Zwar erwähnt Schelling diesen (und auch Ainslee’s) Beitrag (z.B. Schelling, 1984; S. 6) ebenso wie die Arbeit von Smith über Self-Command. In all seinen Veröffentlichungen hat sich Schelling jedoch sehr verhalten (um nicht zu sagen unlauter) dazu geäußert, woher die Idee der zwei zeit-inkonsistenten Selbst, der offensichtliche Anstoß seines Nachsinnens über Self-Command, kommt. Tatsächlich schlägt “the Adam and the smith of systematic economics” (Boulding 1969; S. 1) vor, dass das Problem von Self-Command ein Problem von zwei im Widerstreit stehenden Selbst ist (Smith, 1759, 1982).

Smith nennt diese beiden Selbst den *Man Yesterday* und den *Man Today* und konzeptualisiert den *Man Today* sowohl als eine reale Person, die (zeitlich) nach dem *Man Yesterday* existiert als auch als ein imaginäres Konstrukt im Kopf des *Man Yesterday*. Wie in Meardon & Ortmann (1996; 1996a) ausgeführt, hat Smith den intrapersonalen Konflikt des *Man Today* mit dem *Man Yesterday* als ein 2x2-Basissspiel modelliert, das strukturell vergleichbar mit den Reputations-Spielen ist, die vielfach in der modernen Ökonomie in den verschiedensten Kontexten verwendet werden (z.B. Kreps, 1990). Smith hat also Self-Command als ein Gleichgewicht definiert, das durch interne Reputationseffekte aufrechterhalten wird.

In seinem „Richard T. Ely Vortrag“ vor der American Economic Association sagte Schelling (1984): “Adam Smith, by the way, included a chapter on self-command in this Theory of Moral Sentiments. He meant something different – courage, generosity, and other manly virtues. In my usage, self-command is what you may not need to employ if you already have enough of what Adam Smith meant by it.” (1984; S. 3; siehe auch Schelling, 1980; S. 69). Dieses Statement legt den Verdacht nahe, dass Schelling sein Bestreben, sich selbst durch das Lesen des Kapitels über Self-Command von Smith zu verbessern, nicht realisiert hat.

Für uns besteht kein Zweifel, dass Smith, rund 200 Jahre zuvor, über genau dasselbe Problem nachdachte, welches Schelling analysierte (siehe Anhang). Er benutzte die gleiche Konzeption, die Schelling benutzte, und selbstverständlich hat auch Smith verstanden, dass in dem internen Reputationsspiel das von den beiden Selbst ausgetragen wird, sich das authentische Selbst nicht immer durchsetzen koennen. Der wirkliche Unterschied zwischen Smith’s und Schelling’s Diskussion liegt in der unterschiedlichen Betonung von Lösungsstrategien.

Smith, in seiner brillanten Art, interessierte sich vor allem für die Frage des Ursprungs und der Evolution des Selbst-Bildes – ein Thema, das erst kürzlich neue Aufmerksamkeit erfahren hat (z.B. Palacios-Huerta, 2003; Benabou & Tirole, 2004) – und für gesellschaftliche Sanktionen, die eingesetzt werden, diejenigen zu „bändigen“, deren Self-Command versagt. Schelling dagegen, statt sich auf gesellschaftliche Sanktionen zu verlassen, interessierte sich vor allem für Strategien, die es einem ermöglichen, sich selbst zu verpflichten, so dass ein anderes, widerstrebendes Selbst nicht zum Zuge kommt. Er war sozusagen an Mechanismen für den Umgang mit Situationen interessiert, in denen die internen Reputationsspiele möglicherweise nicht zum erwünschten Ergebnis führen.

Schließlich mag noch ein dritter Grund für den geringeren Einfluss von Schelling’s Arbeit über Self-Command verantwortlich sein: Wie Smith, kämpfte Schelling mit einem genuin schwierigen Problem. Es lässt sich deshalb so schwer formalisieren, weil die Konzeption von uns selbst ein “heterogeneous array of types and circumstances”

(Schelling, 1984; S. 4) beinhaltet, die schwer fassbar (und experimentell kontrollierbar) sind. Zusätzlich gibt es interessante methodische Probleme, die den Fortschritt experimenteller Untersuchungen der Frage zeit-konsistenten Verhaltens verlangsamt haben (z.B. Harrison & Lau, 2005).

4. Fazit

Zeckhauser (1988) meint, dass “Schelling’s work [contributed] fundamental game theory insights to political science, psychology, and sociology long before economists found that conjectural equilibria and commitment difficulties were central to [economics].” (S. 160). Obwohl Zeckenhauser’s Überlegungen aufgrund seiner kollegialen Beziehung zu Schelling nicht völlig neutral zu bewerten sind, halten wir seine Einschätzung – sogar im Hinblick auf die Theorie der Self-Command – für unbestritten. Schelling hat intensiv und mit Genialität darüber nachgedacht, was das authentische Selbst (falls es denn existiert) benötigt, um das andere Selbst zu besiegen, wenn das interne Reputationsspiel sich nicht erfolgsversprechend gestaltet. Dennoch liefert Schelling keine Formalisierung seiner Theorie, mit der Verhalten systematisch analysiert, erklärt und vorhergesagt werden könnte.

Obwohl seine Theorie von Selbst-Command nicht die Aufmerksamkeit seiner früheren Arbeiten erlangte (gleichwohl sie zum Teil darauf aufbaut), ist sie doch stets unterhaltsam, instruktiv und ohne Zweifel sehr bedeutsam, wie auch die kürzlich erschienene Arbeit von Duckworth und Seligman (2005) zeigt. Auch wäre vermutlich dieser Beitrag ohne die Lektüre von Schelling’s Überlegungen zu Self-Command-Strategien nicht geschrieben worden (oder doch zumindest erheblich später).

Anhang

Self-Command in Adam Smith’s *Theory of Moral Sentiments*¹⁰ – Eine Zusammenfassung

“When our passive feelings are almost as sordid and selfish, how comes it that our active principles should often be so generous and so noble? ... It is not the soft power of humanity, it is not the feeble spark of

¹⁰ Dieser Zusammenfassung basiert wesentlich auf Meardon & Ortmann (1996; 1996a).

benevolence which Nature has lighted up in the human heart, that is thus capable of counteracting the strongest impulses of self-love.”
(Smith, 1759, 1982; S. 137)

Nach Adam Smith ist Self-Command die Anwendung von „*active principles*“ um das eigene Verhalten zu zügeln, wenn „*passive feelings*“, die dieses Verhalten auslösen würden, zu stark oder zu übermächtig sind oder dazu führen würden, sich unangemessen zu verhalten. Unter „*passive feelings*“ sind die Gefühle, Leidenschaften, Passionen oder Verhaltenstendenzen zu verstehen, die wir nicht leicht kontrollieren können, die meist als automatisch ablaufend empfunden werden und die unsere Handlungen stark beeinflussen können: Hierzu gehören körperliches Verlangen (wie beispielsweise sexuelle Begierde — “naturally the most furious of all passions” [a.a.O., S. 28]), oder antisoziale Affekte wie Hass und Missgunst in all ihren Facetten. Obwohl sich Smith in seinen Erörterungen über Self-Command mehr auf die Auswirkungen eines Versagens von Self-Command auf Dritte konzentrierte, war die grundsätzliche Frage, nämlich wie das eigene Temperament oder unterschiedliche Verlangen zu kontrollieren sind, die gleiche, mit der sich Schelling befasst hat.

Das Eingangszitat führte Smith wie folgt fort: “It is a stronger power, a more forcible motive, which asserts itself upon such occasions. It is reason, principle, conscience, the inhabitant of the breast, the great judge and arbiter of our conduct. ... “ (Smith, [1759], 1982; S. 137). Es sind also Vernunft, Grundsatz, Gewissen, die die „*active principles*“ ausmachen und die „*passive feelings*“ unter Kontrolle halten. Gegen die meisten unserer Gefühle und Leidenschaften, unserer „*passive feelings*“, können wir nicht viel tun. Sogar die schlimmsten Empfindungen gehören zur menschlichen Natur. Wir können jedoch unsere Handlungen aktiv beeinflussen.

Obwohl wir oft von intensiven Wünschen gedrängt werden – das sexuelle Verlangen zum Beispiel (das uns daran hindern mag, die Steuererklärung fertig zustellen), oder das ganz gewöhnliche Verlangen nach Kuchen, oder Nikotin, oder Fernsehen (was, alleine oder in Kombination, verhindern mag, dass wir die durchtrainierten Luxuskörper werden, die vor Gesundheit nur so strotzen) – sagt uns die Vernunft, dass es nicht in unserem Interesse wäre, ihnen nachzugeben, weil möglicherweise die nicht fertig gestellte Steuererklärung oder unser deformierter Akademikerkörper ein zu hoher Preis sind. In diesen Fällen

müssen wir uns auf unsere „*active principles*“ verlassen, um unseren kurzfristigen Impulsen zu widerstehen.

Soweit, so gut: Wie aber kommen unsere „*active principles*“ zu Stande? Um “the constant necessity, of modeling, or of endeavouring to model not only his outward conduct and behaviour, but, as much as he can, even his inward sentiments and feelings” (Smith 1759, 1982; S. 147) zu modellieren, nimmt Smith an, dass das Problem in zwei unterschiedlichen Selbst besteht, die sich beide verwirklichen wollen: „The Man of Today and the Man of Yesterday“. Der *Man of Yesterday* ist die Person, die durch Leidenschaft entfacht und bereit ist entsprechend zu handeln; der *Man of Today* ist die Person, die anschliessend mit den (unter Umständen negativen) Konsequenzen dieser Handlung konfrontiert ist. Mit den Worten von Smith: „The man of to-day is no longer agitated by the same passions which distracted the man of yesterday: and when the paroxysm of emotion, in the same manner as when the paroxysm of distress, is fairly over we can identify ourselves, as it were, with the ideal man within the breast, ... “ (a.a.O., S. 157f.).

Der *Man Today* existiert sowohl als reale Person (zeitlich) nach dem *Man Yesterday* wie auch als imaginäres Konstrukt im Kopf des *Man Yesterday* (also als „*Man Today*“). Smith macht klar, selbst wenn sich der *Man Yesterday* mitten in den unsozialsten Leidenschaften befindet, ahnt sein eigener Verstand (Smith, 1759, 1982, S. 161) die Konsequenzen, die es haben wird, wenn er ihnen erliegt. Diese Idee des imaginären Konstrukts ermöglicht Meardon & Ortmann (1996) den intertemporären Aspekt des *Man Yesterday* und des *Man Today* weitgehend zu ignorieren und das Self-Command Problem als einstufiges Spiel zu modellieren, bei dem die Teilnehmer simultan agieren.

Meardon & Ortmann (1996) erklären, wie sich Smith die Entstehung der Einsicht dessen, was lobenswert und dessen was verachtenswert ist, vorstellt; sie erklären auch, wie Smith sich den Prozess der Konvergenz des parteiischen („*partial spectator*“) zum unparteiischen Beobachter („*impartial spectator*“=„the ideal man within the breast“, der weiß, was die Gesellschaft und vielleicht einige höhere Instanzen erwarten, der weiss was lobenswert und verachtenswert ist) vorstellt. Auf dieser Grundlage argumentieren sie, dass der intrapersonale Auseinandersetzung zweier Selbst als Spiel modelliert werden

kann in dem der *Man Yesterday* für sein Verhalten zwei Optionen hat — „richtig“ (proper) oder „falsch“ (improper) — und der *Man Today* ebenfalls zwei Optionen hat: die Aktion des *Man Yesterday* hinsichtlich des Aufwandes an emotionalen Kosten entweder mit „weniger“ (routine) oder „mehr“ (real) zu bewerten. Die Autoren belegen, dass sich dieses Szenario in eine 2x2 Basisspielmatrix umsetzen lässt, in der der *Man Yesterday* der Zeilenspieler und der *Man Today* der Spaltenspieler ist. Sie begründen auch im Detail die Payoffs, die die 2x2 Basisspielmatrix zu einer der Auszahlungstabelle werden lassen, die aus Analysen von Qualitäts-Spielen (*quality and effort assurance*) bekannt ist (z.B. Kreps, 1990), und die Elster's scharfsinnige Bemerkung über das Selbstbild („self-image“) als einen internen Reputationseffekt formalisiert (Elster, 1985; S. 265). Mit anderen Worten, Meardon & Ortmann (1996) argumentieren und belegen, dass Smith Self-Command als ein Gleichgewicht konzeptualisiert hat, das durch interne Reputationseffekte aufrechterhalten wird.

Tabelle 1 zeigt die Payoff-Matrix, die Meardon & Ortmann (1996) aus Smith's Arbeiten ableiten:

Tabelle 1: Matrix für das Self-Command Spiel

| | | <i>Man Today</i> | |
|----------------------|----------|------------------|------|
| | | Routine | Real |
| <i>Man Yesterday</i> | Proper | 1,1 | 0,0 |
| | Improper | 2,-1 | 0,0 |

Es ist ein Standardergebnis der spieltheoretischen Literatur (z.B. Kreps, 1990; S. 65ff.), dass das Ergebnis eines Qualitäts- oder Reputations-Spiels entscheidend davon abhängt, wie oft es gespielt wird. Besonders bei halbwegs niedrigen Diskontierungsfaktoren lässt sich leicht demonstrieren, dass ein Nash-Gleichgewicht des unendlich wiederholten Basisspiels Pareto-optimal ist (d.h. in der linken oberen Zelle in der Payoff-Matrix des Self-Command Spiels, lokalisiert ist). Im Gegensatz dazu ist das wahrscheinliche Ergebnis in einem nur einmal gespielten Spiel die untere linke oder rechte Zelle (wobei die rechte Zelle das Gleichgewicht repräsentiert).

Dieses Ergebnis ist intuitiv leicht nachvollziehbar¹¹. In der oberen linken Ecke ist in jeder Runde des unendlich wiederholten Basisspiels die Auszahlung für *ManYesterday* und *Man Today* jeweils „1“. Insbesondere wenn zukünftige Auszahlungen sehr viel Gewicht haben (d.h., wenn der Diskontierungsfaktor verhältnismäßig gering ist), ist die Summe, die *Man Yesterday* und *Man Today* auf diese Weise erreichen können, beträchtlich und ermöglicht es den beiden Selbst, in ihren Interessen überein zu stimmen. In einem nur einmalig gespieltem Basisspiel können diese Interessen aus einem einfachen Grund divergieren: Selbst wenn *Man Yesterday* und *Man Today* sich auf die Pareto-optimalen Auszahlungen in der oberen linken Zelle der Payoff-Matrix einigen, wird für den *Man Yesterday* angesichts des Payoffs in der unteren linken Zelle ein starker Anreiz bestehen, der Versuchung einer Auszahlung von „2“ nachzugeben und damit die Vereinbarung zu brechen. Das würde zwar den *Man Yesterday* viel besser stellen, aber den *Man Today* viel schlechter (beachte dass die Summe der Auszahlungen in der unteren linken Zelle der Payoff-Matrix nicht Pareto-optimal ist). Wahrscheinlich wird aber der *Man Today* den Wechsel der Aktionswahl von „proper“ nach „improper“ vom *Man Yesterday* antizipieren und deshalb seinerseits einen Wechsel der Aktionswahl von „routine“ auf „real“ vollziehen, was beide in der Gesamtsumme sogar noch schlechter stellt als vorher (der „*Man Today*“ erreicht, verglichen mit der unteren linken Zelle, zwar ein etwas besseres Ergebnis, der *Man Yesterday* jedoch ein viel schlechteres), nämlich in der unteren rechten Zelle des Self-Command Spiels als Endzustand. Es ist diese drohende Möglichkeit dass „*Man Today*“ und *Man Yesterday* in der unteren rechten Zelle anstatt in der oberen linken Zelle landen, die letztlich in unendlich wiederholten Spielen ein Pareto-optimales Ergebnis bewirkt.

Allerdings trifft dies nur unter bestimmten Annahmen über Auszahlungen und Diskontierungsfaktoren sowie Rationalität zu. Smith war sich sehr bewusst darüber, dass das Wissen wann und wie die *active principles* einzusetzen sind, nicht bei jeder Person automatisch vorhanden ist. Solches Wissen erfordert eine gewisse Fähigkeit, aktuelle und zukünftige Auszahlungen abzugleichen; eine Fähigkeit, die jedoch erst im Verlauf des Lebens (vielleicht mit großer Anstrengung oder vielen Fällen von Versuch und Irrtum) erworben wird: „A very young child has no self-command.“, schrieb Smith (wie seitdem auch überzeugend experimentell belegt wurde, z.B. durch den genialen „Marshmallow-

¹¹ Für den formalen Beweis siehe Kreps (1990) oder Meardon & Ortmann (1996).

Test“, den Mischel 1968 mit relativ geringen Anreizen durchführte). Wenn aber ein Kind älter wird, “[it] enters into the great school of self-command, it studies to be more and more master of itself, and begins to exercise over its own feelings a discipline which the practice of the longest life is very seldom sufficient to bring to complete perfection.” (Smith, 1759, 1982; S. 145).

Obwohl sich Smith dessen bewusst war, dass Perfektion kaum erreichbar ist, zeichnete er ein Idealbild des standhaften und prinzipientreuen Menschen (der ganz ohne die Art von Tricks, wie sie Schelling als Strategie vorgeschlagen hat, auskommt): “The man of real constancy and firmness.¹², the wise and just man who has been thoroughly bred in the great school of self-command, in the bustle and business of the world, ... maintains this control of his passive feelings upon all occasions; ... He has never dared to forget for one moment the judgement which the impartial spectator would pass upon his sentiments and conduct. He has never dared to suffer from the man from within the breast to be absent one moment from his attention. With the eyes of this great inmate he has always been accustomed to regard whatever relates to himself. This habit has become perfectly familiar to him. He has been in the constant practice, and, indeed, under the constant necessity, of modeling, or of endeavoring to model, not only his outward conduct and behaviour, but, as much as he can, even his inward sentiments and feelings, according to those of this awful and respectable judge. He does not merely affect the sentiments of the impartial spectator. He really adopts them.“ (Smith, 1759, 1982; 146f.).

Nun ist es aber so, dass viele Menschen nicht die Gelegenheit haben, von “the great school of self-command” zu profitieren. Selbst denjenigen, die durch diese Schule gegangen sind, gelingt nicht immer und in jeder Situation ein entsprechendes Verhalten. Ihnen fehlt die echte Standhaftigkeit und Entschlossenheit (“real constancy and firmness”) im Sinne von Smith, weshalb sie Schwierigkeiten haben, die relevanten Einschätzungen angemessen zu vollziehen. Eben diese Menschen brauchen Methoden, sich selbst zu überlisten und sie können bestimmt von Schelling’s Sammlung von

¹² In seinem „Richard T. Ely“ Vortrag vor der American Economic Association hat Schelling (1984, S. 4) Smith für seine, aus heutiger Sicht politisch nicht korrekten Formulierung angegriffen. Es scheint uns eine interessante Frage zu sein, was anstößiger ist: Jemand der im 18. Jahrhundert darüber spricht, was er als geschlechtsspezifisches Verhalten betrachtet oder jemand der, im Jahre 1984, anderer Leute Ideen als seine reklamiert.

Selbstverpflichtungsstrategien und Tricks profitieren, weil diese Strategien, anstatt sich auf reputative Selbstverpflichtungen zu verlassen, intelligente Mechanismen bereitstellen, die solche reputative Selbstverpflichtungen¹³ unnötig machen.

Die Zusammenfassung von Smith's Diskussion über Self-Command sollte verdeutlichen, dass Schelling nicht sehr mitteilsam und offen war, woher die Idee der zwei zeit-konsistenten Selbst (klarer Ausgangspunkt seiner Überlegungen zu Self-Command) ihren Ursprung hat. Ohne Zweifel – Smith sprach über exakt dasselbe Problem wie Schelling und er benutzte dieselbe Konzeptualisierung, die Schelling benutzte.

¹³ Palacios-Huerta (2003) bezeichnet solche reputativen Selbstverpflichtungen als "*internal commitment technologies*" (S. 249), was ein brauchbares Kuerzel ist, um sie von Schellings's „external commitment technologies“ zu unterscheiden.

Literatur

- Ainslee, George (1975): Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control, in: *Psychological Bulletin* 82.4 (July), S. 463–96.
- Anonymous (2005): Interview: Thomas Schelling, in: *Region Focus* (Spring), www.richmondfed.org, S. 36–41.
- Benabou, Roland und Jean Tirole (2004): Willpower and Personal Rules, in: *Journal of Political Economy* 112.4, S. 848–886.
- Boulding, Kenneth E. (1969): Economics As a Moral Science, in: *American Economic Review* 59.1, S. 1–12.
- Duckworth, Angela L. und Martin E.P. Seligman (2005): Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents, in: *Psychological Science* 16, S. 939–944.
- Harrison, Glenn W. und Morten I. Lau (2005): Is the Evidence for Hyperbolic Discounting in Humans Just An Experimental Artefact? In: *Brain and Behavioral Sciences* 28, S. 657.
- Kreps, David M. (1990): *Game Theory and Economic Modelling*. Oxford.
- LeDoux, Joseph (2002): *Synaptic Self. How Our Brains Become Who We Are*. New York.
- Meardon, Stephen und Andreas Ortmann (1996): Self-command in Adam Smith's *Theory of Moral Sentiments*: A Game-theoretic Reinterpretation, in: *Rationality and Society*. 8.1 (February), S. 57–80.
- (1996a): Yes, Adam Smith Was an Economist (A Very Modern One Indeed), in: *Rationality and Society*. 8.3 (August), S. 348–52.
- Mischel, Walter (1968), *Personality and Assessment*. New Work.
- Palacios-Huerta, Ignacio (2003): Time-Inconsistent Preferences in Adam Smith and David Hume, in: *History of Political Economy* 35.2 (Summer), S. 241–58.
- Schelling, Thomas (1960), *The Strategy of Conflict*. Cambridge.
- (1978): *Micromotives and Macrobehavior*. Cambridge
- (1980): The intimate contest of self-command, in: *The Public Interest* 60 (Summer), S. 94–118. (Reprinted in Schelling 1984a; chapter 3, S. 57–82).
- (1982): Ethics, Law, and the Exercise of Self-Command. (Tanner Lecture at the University of Michigan). (Reprinted in Schelling 1984a; chapter 4, S. 83–112).
- (1984): Self-command in practice, policy, and in a theory of rational choice, in: *The AER, Papers and proceedings*. 74.2, S. 1–11.
- (1984a): *Choice and Consequence. Perspectives of an errant economist*. Cambridge.
- (1992): Addictive Drugs: The Cigarette Experience, in: *Science* 255 (Jan 24), S. 430–33.

— (2006): *Strategies of Commitment and Other Essays*. Cambridge

Smith, Adam (1759, 1982): *The Theory of Moral Sentiments*. D.D. Raphael and A.L. Macfie (Hrsg.). Indianapolis.

Strotz, Robert H. (1956): Myopia and Inconsistency in Dynamic Utility Maximization, in: *Review of Economic Studies* ??, S. 165–80.

Zeckhauser, Richard (1989): Reflections on Thomas Schelling, in: *Journal of Economic Perspectives* 3.2 (Spring), S. 153–64.